An extension of the use of biodata for managerial selection

Melanie Mitchell

University of Northumbria at Newcastle

A thesis submitted in partial requirement

for the degree of Doctor of Philosophy

December 2001

Acknowledgements

First and foremost thanks must go to my supervisor, Dr Larry Wilkinson for his boundless energy, enthusiasm and for making it look easy. I would also like to express my extreme gratitude to my supervisor Nigel van Zwanenberg for his unending supply of inspirational ideas and philosophical deliberation throughout the project. I would like to thank Newcastle Business School at the University of Northumbria at Newcastle for the studentship that enabled me to study for this degree on a full-time basis.

Sincere thanks must to the organisations involved and all of the volunteers without which this research would not have possible. I would particularly like to thank Carol McCletchie for her support.

Further gratitude is due to my parents for their unmoving belief and support from the beginning and to numerous friends and family members, human and otherwise for providing ears and smiles as required. A special mention must go to my sisters and to Reece for providing laughs. A huge thank you to my office-mates, to Ruth for providing entertainment, amusement and for showing me the sights, to Lana for her wisdom and good advice and to Amanda for understanding.

Thanks to Mark for everything. This thesis is dedicated to you with love.

An extension of the use of biodata for managerial selection

Abstract

The aim of the research was to investigate whether a construct-oriented approach to biodata modelling provides incremental validity over and above other instruments currently employed in the selection of managers. This aim was explored through the development of construct oriented biodata analogues of the constructs of critical thinking ability, extroversion and neuroticism. These models were developed on a pilot sample of 'potential managerial candidates'. The pilot analogue models demonstrated impressive levels of construct validity and the biodata instrument was then validated in a concurrent study based upon managerial job incumbents. Supervisor ratings of performance and a career progress variable provided subjective and objective indicators of managerial performance.

Although the psychological constructs of critical thinking ability, extroversion, and neuroticism did not significantly predict either outcome, further analysis of supervisor ratings revealed that perceived conscientiousness and energy contribute much of the variance associated with overall performance ratings, suggesting the likelihood of halo error in the ratings and offering grounds for a social psychological explanation of the results relating to this criterion.

Regression analyses revealed that biodata analogue models of critical thinking ability, extroversion and neuroticism demonstrate incremental validity of construct-oriented biodata analogue models over traditional psychometric measures of these constructs. Construct-oriented biographical life history analogues may add considerable utility when used in the pre-selection stage of managerial recruitment and selection.

Contents

Tables

Figures

Glossary of terms

Biographical data / biodata: this term refers to any information incorporating or relating to events or experience from the background life experience of an individual, such as educational and employment history, and family background information.

Concurrent validity study: this approach to assessing the value of a measure or predictor involves collecting 'predictor' (for example, the selection method) and 'criterion' (e.g. performance ratings) information from existing employees within the same period of time and correlating the two to obtain the validity coefficient.

Construct-oriented biodata: biographical life history items designed to capture variance associated with specified psychological constructs. Such constructs have normally been selected on the basis of a perceived theoretical association between the construct as a predictor of a chosen criterion, or performance output.

Convergent validity: the extent to which constructs that should theoretically be related or show correspondence, do so in the form of significant correlation coefficients. This is necessary but not sufficient for construct validity.

Correlation coefficient: a number between –1 and +1 representing the relatedness of any two variables, where (+ or -) 1 represents a perfect correlation and 0 represents no relation between the variables.

Cross-validation: a repeated validation study of models derived in an initial study upon a second sample. The procedure ascertains the 'robustness' of validity results as indicated by small amounts of shrinkage in the validity, affirming the result as true and not dependent upon randomisation on chance or artefact.

Discriminant validity: the extent to which a validity study provides evidence that a construct is not correspondent to or related to a construct that should be theoretically distinct. This is necessary but not sufficient for construct validity. Convergent and discriminant validity together, are required in order to be confident that a measure has construct validity.

Incremental validity: any additional validity gained through the inclusion of an additional measure or predictor. In this project this is studied with reference to the improvement of the validity of biographical data analogues of psychometric measures in comparison to the original measures of psychometric constructs as predictors of a managerial performance criterion.

Predictive validity study: this approach necessitates the collection of a predictor (e.g. scores on a selection method) and its correlation with a criterion collected at a later date (e.g. performance ratings) in order to observe how well the selection method predicts job performance in the future.

Chapter 1 Introduction and aims of the research

1.1 Rationale and justification for the research

This thesis outlines a search for the prediction of behaviour. If behaviour is conceptualised as the interaction between person and situation, an approach involving the integration of traditional trait personality theory and social psychological perspectives is required. The Eysenckian model of personality wherein the pattern of personality responses or preferences is embedded within temperament and biological makeup and is expressed through superordinate traits, general preferences, general habits and specific behaviours, offers scope for the assessment and understanding of personality (Eysenck & Eysenck, 1963). The simplicity of this account is compelling. The possibility of capturing the essence of character with reference to two broad traits of extroversion and emotional stability, and to represent this character using a standardised psychometric test scoring procedure is intriguing. However, it is important to appreciate that prediction of the second half of the equation, the interaction between personality 'X' and the world, is less possible. In the social psychological literature, the use of attitude questionnaires to predict behaviour towards attitude objects has notoriously produced low to moderate correlations. A purely personality-based or situation-based focus provides a narrow-band explanation and approach to the prediction and understanding of behaviour. This is borne out by the low to moderate correlations between scores on psychometric measures of personality constructs, and behaviours that are said to represent or embody that construct. Personality measures typically collect 'signs' or preferences toward behaviours rather than actual 'samples' or recollection of actual behaviour. As the best predictor of future behaviour is the way that the individual has behaved in the past (Wernimont & Campbell,

1968), the prediction of behaviour requires a marriage of personality or individual differences accounts of psychology with a broader social psychological or micro-sociological focus. Behaviour is conceptualised as the reciprocal interaction between the individual and the surrounding environment. This interactionist approach is encapsulated in the present research through the extension of current developments in the growing field of construct-oriented biographical life history data. The sum of past life experience or 'biographical data' should encapsulate the personality or traits as well as the range of learning experiences and history of reinforcement that express the interaction of inbuilt traits within the dynamic environment.

The issue of employee selection requires the prediction of performance-related behaviour within the working context. The typical approach to the development and validation of selection methods involves the design of (for example) an interview technique, aptitude test or assessment centre exercise, the scoring of the procedure, and then the correlation of this score with one obtained from an index of performance. This validation process offers insights into the structure of the work performed and may necessitate the development of a person specification based upon the profile of a successful worker. The utility of the selection technique is embedded within the reliability and validity of the performance criterion against which it has been validated (see chapter 5 for a discussion of issues relating to managerial performance appraisal).

The development of approaches to managerial selection is especially interesting as the work of the manager may be relatively autonomous, hence managerial job performance appraisal requires skilled consideration. As a result individual differences such as

cognitive ability and personality may be considered to have significance in mediating performance (see chapter two for a discussion of managerial selection).

Figure 1.1 illustrates the central issues for consideration in the development of an approach to managerial selection. This diagram is intended to reveal the over-lap of areas and the impact of individual differences approaches upon ach of the three key strands of this thesis.

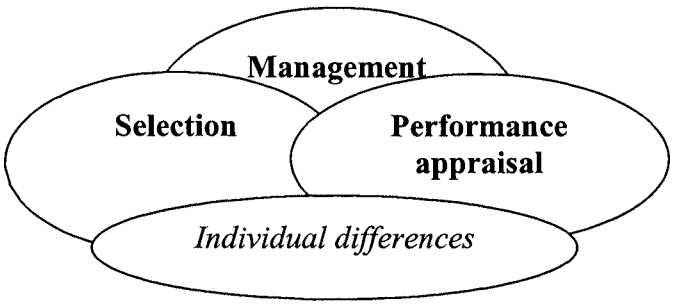


Figure 1.1 Fundamental themes within the thesis

1.1.1 Individual differences as predictors of managerial success

With regard to intellectual skill, the validity of cognitive ability measures for selection of a wide range of occupations is well documented (for example, Ghiselli, 1966, 1971, 1973).

Evaluation of the efficacy of measures of personality for managerial selection necessitates brief consideration of empirical validity, the extent to which the theory offers a parsimonious explanations of the behaviour and results observed, the comprehensiveness of the theory, the internal consistency of the theory, and strategy for personality measurement and change.

Kline (1992) conceptualises the psychometric model as the assumption that all that an individual does may be predicted in terms of characteristics or attributes, abilities, motivation and situational influences. Methods of employing the psychometric model in selection could involve the selection of applicants sharing personality characteristics with existing successful staff as in a standard concurrent validity approach, calculation of the regression equations between personality profiles and job success or the selection of candidates with profiles that may complement existing staff. Kline (1992) asserts that personality is implicated in job satisfaction and success and that psychometrics can offer validity and reliability at greater levels than more subjective selection methods such as interviews and references.

The Eysenckian model (Eysenck & Eysenck, 1963) of personality represents a strand of the dispositional approach to personality. The assumptions of this strategy may be summarised as follows: dispositions are stable and enduring within the individual, they have some consistency and generality for each person and the differences we observe between individuals arise from differences in the strength and amount of dispositions each person has. This approach is also known as the 'nomothetic' approach to personality as it is focussed upon the degree to which we are similar or dissimilar from one another rather than an idiographic or 'the unique individual' approach (Phares & Chaplin, 1997). Personalities are 'disposed' to act in certain ways and this should be consistent across time and situations. These dispositions are referred to as traits and are distinct from temporary states (such as fatigue, under stress etc.). The nomothetic approach permits standardised testing of personality traits on a self-report or peer and observation report basis. Such personality inventories aim to capture the psychological construct of interest but will also collect background variance or 'noise' to some extent. Limitations of the dispositional

strategy include the limited explanations of the basis of traits prior to Eysenck's theory of cortical arousal as the basis for extroversion and autonomic reactivity as the basis of emotional stability. Descriptions of the trait characteristics however, do not provide specific prediction of the likelihood of the observation of a specific trait in a specific situation.

By contrast, Mischel (1977, 1981) proposes an integrative model of personality and situational interaction. This approach challenges the assumptions of traditional dispositional approaches that may assume that test responses are a '***sign***' of personality. Personality assessment (and the prediction of behaviour) should be based upon ***'samples'*** rather than signs of behaviour (Wernimont & Campbell, 1968). Mischel (1977) coined the phrase 'personality coefficient' to refer to a ceiling of +0.30 when self-report measures of personality are correlated with another measure of the same dimension and behaviour.

Mischel outlines 'person variables' that may be considered to be central to an integrated understanding of personality. Person variables include behavioural and cognitive competencies, encoding strategies and personal constructs about the nature of the world, expectations about the consequences of behaviour and situations, value judgements on situation and behaviours, and self-regulatory systems of rewards and punishments aligned with implicit personal goals and externally influenced rules and norms for behaviour.

1.1.2 Biographical life history data and construct-oriented biodata

Biographical data or 'biodata' refers to any information about personal life history and experience. The design of biodata and the nature of the items may vary considerably depending on the choices of the organisation when developing the questionnaire. Biodata is most often collected using a self-report questionnaire format with multiple-choice answers.

The use of biographical data inventories and questionnaires at the pre-selection stage appears to be increasing (Anderson & Shackleton, 1986; Robertson and Makin, 1986) However, Harvey-Cook (2000) reports the more common picture of using heuristic judgments based upon application form information of previous performance in examinations, work experience and hobbies and interests to make decisions regarding suitability for employment.

Biodata represents one way of obtaining information about past behaviour for the purpose of making predictions about future performance or behaviour. Biodata questionnaires obtain information in a consistent way for all candidates from items that are known to be related to job performance (Owens, 1976). Information is scored in a standardized way for all candidates. This technique of obtaining information can address some of the concerns about error and unreliability often associated with less standardized methods of selection, the unstructured interview for example (Conway, Jako, & Goodman, 1995).

To date, a number of studies have explored the possibility of developing construct-oriented biodata. Mael (1991) argues that the inclusion of ability, temperament, and

motivational influences in biodata inventories can result in the higher predictive validity of biodata in comparison to the original personality and interest measures for military samples. This may be attributable to the fact that such biodata attempts "to capture the personal identity and the range of social identities" of the individual; in contrast, personality measures generally focus upon the personal identity or 'level of abstraction' only (Turner, 1985). This approach has been developed on military and civilian occupational samples (Kilcullen, White, Mumford & Mack, 1993; Wilkinson, 1995; Mumford, Costanza, Connelly, & Johnson, 1996).

McCrae & Costa (1988) argue that as parents influence the development of personality, so biographical life history items should relate to personality. Personality in childhood should also influence the situations that a child may find herself in and the experiences she goes through providing the material for later biodata responses. In this way, biodata can offer greater insight into the reciprocal relationship between personality and environment than traditional personality inventories. Indeed, Hough (1989) argues that there is considerable overlap between biodata and personality items though biodata typically captures a broader domain than personality items (Mumford, Snell & Reiter-Palmon, 1994; Nickels, 1994).

The current research provides insight into the incremental validity that may be demonstrated by biodata designed to capture psychometric constructs associated with job performance ('construct-oriented') in the context of managerial selection.

1.1.3 The problem of the performance criterion

In the present research, the development of a biographical information questionnaire that could offer an off-the shelf managerial pre-selection tool necessitates the utilisation of a generalisable performance criterion. Criticisms of traditional biodata include the observation that the scales must be newly validated in each organisation. The development of the current pre-selection instrument requires a criterion that is general across managerial levels and function, generalisable across organisational and departmental boundaries, and worker-oriented in order to reflect changes in the nature of the work of the manager without job change having a deleterious effect on scale validity. Limitations associated with traditional task-based approaches to performance measurement and appraisal includes the measurement of non-observable behaviour, a lack of attention to organisational goals and cultural influence, and rating contamination (Campbell, 1994). A 'dynamic criterion' focus could delineate specific characteristics or abilities that are good predictors of performance over time, and across job requirement changes (Ghiselli, 1956). In addition to temporal changes in the job, the scope for autonomic action or discretion in behaviour choices, may amplify the criterion problem when applied to managerial work (e.g. Borman, 1991). The possibility that two managers in the same job may be equally effective for different reasons requires consideration. Performance can thus be conceptualised as multidimensional, therefore criteria should be multidimensional (Borman, 1991).

Fine and Wiley (1955) present Functional Job Analysis as a system for classification of all occupational functions into interaction between the worker and people, data, and things. This model will be adapted to observe managerial behaviours across organisational and

hierarchical boundaries to form part of the performance appraisal required in this research.

1.2 Research problem

This research project focuses upon the application of a construct-oriented biodata instrument in an occupational context. Biodata analogue models of three psychological constructs (critical thinking ability, extroversion and neuroticism) will be used to predict managerial job performance. The incremental validity of the biodata will be assessed in comparison to the predictive validity of the original psychometric tests of these constructs. The development of an instrument for managerial selection requires consideration of the nature of managerial work. Demonstration of the reliability, validity and utility of such a tool also required the development of an appropriate criterion to be used for the establishment of criterion (predictive or concurrent) validity.

The general research question addressed in this thesis is as follows:

> ***Do construct-oriented biodata analogue models of psychometric constructs have incremental validity over traditional measures of these constructs in the prediction of managerial job performance?***

The aim of this project therefore, is to validate and extend previous research into biodata and to identify any incremental validity that the new biodata models may possess.

The biodata under current development is designed to focus upon the characteristics of the manager and, should it prove to demonstrate criterion-related validity, will offer a tool for the selection of managers based upon characteristics often associated with managerial success (i.e. critical thinking ability (a measure of reasoning and logical thinking), and

extroversion), in addition to a preliminary investigation of the potential moderating effects of neuroticism (Eysenck & Eysenck, 1963) upon performance.

1.3 Research strategy and summary of methodology

The current research utilises a construct-oriented approach to biodata development. This refers to a focus during the item development stage upon psychological constructs believed to underpin or predict psychological constructs associated with typical rather than maximum performance (as assessed by the Watson Glaser Critical Thinking Ability and the Eysenck Personality Inventory measures of extroversion and neuroticism). Construct-oriented biodata may therefore provide a better indication of performance over sustained periods of time than 'snapshot' approaches to measuring ability. In terms of personality, the biodata questionnaire measures preferences for work (team versus individual, autonomous / independent versus directed and other-led), team roles, hobbies and self-perceptions. Previous research documents the decreased likelihood of biodata being subject to response distortion in the way that traditional psychometric measures of personality characteristics may be (e.g. Mael & Hirsch, 1993).

The present research extends the work of Wilkinson to observe the potential incremental validity of a construct-oriented biodata approach in a sample of managers. As such, this thesis presents the results of the application of construct-oriented biodata scales in concurrent and incremental validity studies. The experimental design reflects the standard occupational psychology paradigm, wherein a criterion consisting of an index of job performance will be devised so that the predictor may be calibrated against it.

Incremental validity of inductively developed biodata measures is expected to occur in the current research because of the nature of biodata as a 'sample' of previous behaviour in comparison to traditional psychometric test measures or 'signs' of behaviour based upon individual difference constructs. Additional validity will be assessed by performing regression analyses on biodata analogue models with original psychometric test measures, against performance criteria. In this instance, the performance criteria involved are supervisor ratings of overall performance and an objective career progress variable.
This research has two phases:

- A preliminary study focused upon the development of additional biodata items, new analogue models and the cross validation of Wilkinson (1995).

- The incremental validity study (applying a concurrent validity approach) involving a sample of practising managers.

1.3.1 Preliminary study

The purpose of the preliminary investigation was to demonstrate the validity of the construct–oriented approach to the development of biodata and to explore the generalisability of Wilkinson's models. This study also offered the opportunity to pilot the extended questionnaire in preparation for the incremental validity study with the managerial job incumbents.

The preliminary study involved the following phases:

1 Extension of Wilkinson's (1995) questionnaire in order to include items that may predict neuroticism.

2 Construction of weighted item and unit-weighted (rational) biodata analogue models using a sample similar to that of Wilkinson's 'potential managers'.

3 Explore the potential cross validity of Wilkinson's (1995) original analogue models of critical thinking ability and extroversion.

Analysis began with the construction of analogue models of the psychometric constructs under investigation. Unit-weighted and rational models were developed for the current samples using the same procedure described in the preliminary study methodology (section 4.3 above). The criterion-related validity of the biographical analogues and the original psychometric measures was examined concurrently by performing standard correlational analyses. The incremental validity of the biodata analogue models was ascertained by conducting linear regression analysis in order to observe additional variance accounted for by the analogue models (as in Hunter & Schmidt, 1999).

This preliminary study explores the construct and cross validity of construct-oriented analogue models. In order to form conclusions about the robustness of this approach for biodata development, reliability analysis and construct validity analysis were performed in addition to a standard cross validation exercise (Cronbach, 1990) on the models developed by Wilkinson (1995).

1.3.2 Central investigation

A database of medium to large public and private sector organisations in the region was constructed. The personnel director of each organisation was then contacted with a research brief (see Appendix III). A presentation was delivered to all potential volunteers and the nature of the research project issues of confidentiality and the voluntary nature of participation were discussed. The decision to focus the research on a small number of organisations, including one large sample from a public sector organisation and a small sample from two private sector organisations, represents a compromise from the need to observe the generalisability of the instruments while making data collection practical.

Data collection progressed by conducting psychometric testing sessions with groups of 25 participants. Each managerial job incumbent completed the Eysenck Personality Inventory (obtaining a score on Extroversion and Neuroticism scales) and the Watson Glaser Critical Thinking Appraisal in addition to the biographical data questionnaire. Testing sessions lasted for approximately one and a half hours and all participants received written feedback on the psychometric test scores via an anonymous code matched to that on the performance data. All candidates retained absolute anonymity from the researcher although feedback interviews and telephone counselling were made freely available to all participants wishing to obtain more in-depth understanding of test scores.

The managers were deliberately chosen to represent a range of managerial functions and levels. A sample of 135 managers undertook testing as managerial job incumbents for the concurrent validity stage of the research. A further sample of 33 senior managers completed a performance appraisal proforma for each job incumbent. The next stage of the

research involved the exploration of the concurrent validity of the biodata analogue models using an index of job performance of managers as the criterion. A linear regression analysis reveals additional validity contributed by additional variance captured by biodata analogues in comparison with the traditional psychometric measures of the constructs.

The approaches adopted for the assessment of validity in this investigation are those typically adopted in studies assessing selection techniques (e.g. see Cascio, 1991 for a review). Put simply, a typical validation procedure has been adopted and the biodata models are subjected to scrutiny of construct validity as well criterion related validity and reliability. Discriminant validity is one facet of the psychometric approach that has not been adopted for biodata model development. Items in the biodata models are expected to demonstrate very high inter-correlations and have not been removed on this basis. It is expected that biodata items may capture variance in addition to the trait they have been designed to capture and it may be this extraneous variance that improves prediction of job performance.

1.4 Outline of the thesis

Chapter two presents an introduction to managerial selection with discussion of issues relating to popularity, validity, reliability and the utility of standard approaches to managerial selection.

Chapter three presents an historical review of approaches to the development of biodata models with discussion on the merits and criticisms of each. Alternative approaches are discussed, as is Wilkinson's (1995) construct-oriented approach. Issues of validity and reliability, including socially desirable responding and adverse impact are raised in relation to each of the approaches.

Chapter four discusses the nature of the criterion necessary for the validation of selection instruments. In order to appraise managerial performance the literature discussing the nature of managerial work in terms of actual behaviour and the roles managers engage in is reviewed (in section 3.1). The role of individual differences such as personality variables and organisational citizenship behaviours are discussed and a generalisable approach to managerial appraisal is put forward, based upon Fine and Wiley's (1974) Functional Job Analysis.

Chapter five outlines several of the standard approaches to the collection of performance information including the collection of personnel data and subjective rating systems with particular reference to potential sources of rater error.

Chapter six presents the methodology for phase one of the research. This chapter details the steps taken in the cross-validation study of Wilkinson's original biodata analogue models of the constructs of critical thinking ability and extroversion. The development and

selection of items to extend Wilkinson's (1995) original questionnaire is also debated. The characteristics of a developmental sample of 'potential managers' are outlined and the development of further analogue models for critical thinking, extroversion and neuroticism is described.

Chapter seven presents the methodology of the incremental validity study. This section explains the paradigm adopted by this research and the traditional approach to validation represented throughout the majority of the occupational psychology literature. The construction of the biodata models is discussed in this section. Descriptive data about the sample of managerial job incumbents and the senior managers who completed the performance appraisal proforma is presented in this section.

The results of the incremental validity investigation are presented in chapter eight. This section illustrates the incremental validity of the biodata analogue models predicting the constructs of critical thinking ability, extroversion and neuroticism. Chapter eight presents the results from additional analyses upon the sub-component of total performance ratings and correlates of 'career progress'.

Chapter nine presents additional findings with a focus upon the sub-components of managerial performance ratings with reference to ad-hoc, empirical biodata analogue development for the prediction of perceived conscientiousness and energy.
Chapter ten offers a discussion of the results in relation to the previous literature and the limitations of the current research. Suggestions for further research are outlined and the implications for the present findings are put forward.

Chapter eleven outlines the final conclusions of this research with discussion of the limitations associated with the design and methodology.

Chapter 2 Issues in managerial selection

This section presents an outline of current research findings on the most popular approaches and techniques used for managerial selection so as to provide a background against which to assess the role of biodata in this field.

2.1 Approaches in managerial selection

Bergwerk (1991) asserts that managerial selection often boils down to recruitment of individuals considered 'safe', i.e. they have the right sort of experience and a proven track record, although criteria may often be vague. This method has varying success rates and even the briefest glance at standard utility models of selection illustrates the substantial costs involved in selection mistakes. This can be especially true for managerial staff.

Various hypotheses have emerged explaining the continued popularity of the traditional interview. Among the most persuasive of these, Arvey and Campion (1982) discuss the importance to the organisation of selecting candidates that "fit-in" with organisational culture and existing staff. The interview very likely serves purposes other than simple selection of candidates. It may function as an opportunity to 'sell' the company to candidates, and to exchange information about the organisation (Herriot, 1989). As such, the selection process must be viewed as a dynamic and interactive process, due consideration will be given to applicant reactions to selection procedures.

2.2 The validity of current methods of managerial selection

Perhaps the first investigations into the validity of selection methods date back to Munsterberg (1913, cited in Schmitt and Landy, 1993). Munsterberg argued that test development could be followed by criterion-related studies to achieve a validity measure that represents the predictive validity of any selection technique. In standard validity studies, validity is assessed through a correlational analysis between the selection index or 'predictor' and a performance outcome or 'criterion'(Cronbach, 1990). Schmidt and Landy (1993) defined this process as one of continual hypothesis development and testing following a job analysis procedure to define the performance domains for investigation (e.g. the Position Analysis Questionnaire, McCormick, Jeanneret, & Mecham, 1972.)

Legal guidelines on selection procedures require comprehensive documented validity for all procedures used for making employment decisions. The Institute of Personal Development guide on recruitment (1996) suggests, "The selection process should consist only of properly validated tools and techniques." (IPD, 1996:6). The American Psychological Association (APA) Standards (1985) commission recognize several strategies for evaluating the validity of selection methods. The majority of the research conducted into the efficacy of selection tools for managerial selection and promotion gathers criterion-related evidence of validity through predictive or concurrent validity analysis. Ascertaining the criterion-related validity of a selection technique involves obtaining a correlation coefficient of the relationship between a predictor of performance (or selection technique, for example, assessment centre rating) and a criterion measure (for example, supervisor ratings of performance). In a concurrent validity study, this relationship will be observed by recording the predictor and criterion variables at

approximately the same time. In a predictive validity study, performance on the predictor measure will be compared with performance on a criterion measure obtained after selection/promotion (Guion, 1976; Cascio, 1991). Guion (1976) asserts that "predictive validation is the most appropriate research paradigm for employment testing" but a "predictive validity coefficient is not necessarily a good appraisal of the appropriateness of a test as a selection instrument" (Guion, 1976:789). Disadvantages of the concurrent validity approach include the inevitable restriction of range occurring where volunteers in validation research are necessarily employees of the organisation and have therefore been subject to prior selection. However, the concurrent validation approach offers the researcher numerous advantages including the convenience of obtaining criterion data within a relatively short time frame.

If the selection procedure focuses upon work products rather than 'ability or potential', content-related validity evidence may be obtained from test performance and work sample tests. The content validity ratio represents the extent to which capability to function in a job performance domain and performance in the test under investigation are perceived to overlap. This method is advantageous in that it emphasizes job analysis procedures and improved performance and/or behaviour sampling.

Methodological issues to be considered in the validation of any selection technique include the representativeness of the sample, sample size and the possibility of range restriction in the sample. Criterion contamination must similarly be considered before validity coefficients are acceptable. Bias at the level of the criterion is considered in detail in Chapter five.

Enhanced cost-effectiveness of any method of selection may be summarised as a reduction in the cost and duration of selection methods procedures and/or the introduction of multiple stages in the selection process to reduce the number of candidates present at each stage. The latter recommendation largely represents the final common pathway of selection procedures. For the main part, candidates applying for a position will submit a covering letter with curriculum vitae or an application form in response to a job advertisement. A few will be invited for interview, less still will be invited for a second interview and finally, an accepted candidate will provide references sent out from previous employers (Smith and Robertson, 1986).

In recent years much of the validity research has involved synthetic analysis or generation of predicted validity coefficients across a range of organisations and samples. This approach may also involve adopting a meta-analytical approach to published validity studies. Meta-analysis may be useful in illustrating the predictive validity of each method across different situation and organisational boundaries. The situational specificity of validity has direct relevance for the development of scientific theory and also has practical implications. If a validity level is consistent, the relationship between constructs is not likely to be dependent upon features of the job/organisation/situation. This is important in assessing the cost-utility ratio of selection procedures and is of special consideration when applying methods to small sample sizes such as those found in senior, and sometimes middle level management.

Differential validity or 'adverse impact' by age, sex, race, disability, or sexuality must be considered and investigated. Job performance must be considered in relation to test/procedure performance because tests cannot be said to be discriminating unfairly if test performance continues to predict actual job performance well (Cascio, 1991). Overall,

selection methods "should be fair and valid and recruiters should be able to demonstrate the rationale for the selection tools used." (IPD, 1996:7).

Robertson and Makin (1986, 1993) report survey results of managerial selection techniques used in one hundred and eight British organisations recruiting between one and 995 managers annually. They use their results to dichotomise selection methods into mainstream techniques (the interview, assessment centres, psychometric tests and work sample tests) and non-standard methods (biodata, graphology, astrology). Non-standard techniques that do not demonstrate replicable validity (such as graphology and astrology) will not be discussed here.

The functions of employee selection have been defined as incorporating an opportunity for information gathering; an opportunity for the prediction of future performance and work behaviour; a decision making exercise between candidates and organisations; and giving candidates the opportunity to learn more about an organisation (Roe & Grueter, 1989). The value of a combinative model such as this lies in the view of selection as a two-way interactive process and in highlighting the need to consider the impact of methods upon the candidate as well as the validity, reliability, utility and practicality of selection methods. This chapter presents a review of the most popular approaches to managerial selection and considers the validity and reliability of each method in turn. Biographical data will be considered in detail in a later section (chapter 3).

In conclusion, the assertion of Cascio (1991) that managerial selection requires both the adequate description of the components of executive and managerial success in behavioural terms and the development of behaviourally based predictor measures for the

prediction of success appears to leave the majority of current practice behind. The popular existing selection methods most frequently include the "classic trio" incorporating the application form, interview, and references and the apparent neglect of more reliable or valid measures. This reveals a large gap between research findings and practice (Cook, 1993, 1998).

Despite the lack of consensus and the continuing controversy about what constitutes managerial effectiveness, it appears that this goal may become clearer as the technologies of job analysis and person specification provide more reliable insights into the world of managerial work. Such interventions must necessarily take into account the proposition of Robertson and Makin (1986) that managerial efficacy can be related to, and described in terms of functions of the interaction of the manager and the working environment (Robertson and Makin 1986: 52).

2.2.1 The selection interview

Robertson and Makin (1986, 1993) report that interviews were used almost universally (only one percent of people surveyed report never using them) in their sample of 108 British organisations. The number of interviews and number of interviewers used differs substantially between companies. A clear preference for using more than one interview emerged (approximately two thirds of organisations always used two or more interviews for managerial selection, over a third of organisations reported the consistent use of two or more interviewers). Robertson and Makin (1986, 1993) conclude that despite decades of unimpressive validation data, the interview remains popular and use does not appear to be diminishing. These findings support the earlier work of Kingston (1971) and Beason and

Belt (1976) reporting popularity statistics in the region of between 82 and 88 per cent of organisations for use of interviews. Smith and George (1994) also discuss the continued popularity of the interview. This popularity may at least partially be attributed to the usefulness of the interview in affording opportunities to assess the suitability of the candidate in terms of "fit" with the organisation and the opportunity to clarify and exchange information on the part of candidate and organisation (e.g. Shackleton & Newell, 1989). For example, Herriot (1989) argues that the popularity of the interview may derive from the recognition of the opportunity for the social processes of feedback, accommodation and negotiation in the two-way process of selection. Herriot (1989) concludes that validity problems associated with the interview could be outweighed by its versatility.

Robertson and Iles (1988) postulate that problems associated with the validation of interview methods result primarily from the subjective nature of the interaction. Problems of discrimination and the reliability of the interview relate to the expectations, prejudices and cognitive and affectional biases of the interviewer (Arvey, 1979). Similarly, Schmidt and Noe (1986) discuss the influences of perceived role congruence and the expectation that managerial work is assumed to incorporate the effectiveness of essentially masculine characteristics and the effect this may have upon female recruitment. Kinicki and Lockwood (1985) describe the influence that interviewer impressions of and attraction to candidates can have on the assessment of skills and suitability for hire. Robertson and Iles (1988) also conclude that interviewers typically rely upon impressionistic rather than verifiable information in judging candidates.

In summary, the evidence for the validity of the traditional, unstructured interview suggests that it has low predictive validity and may be unreliable (e.g. Webster, 1964; Ulrich and Trumbo, 1965; Arvey and Campion, 1982). The un-corrected validity coefficient of the traditional (unstructured) interview has been reported to average around 0.2-0.35 (Cortina, Goldstein, Payne, Davison, & Gilliland, 2000:328). Table 2.1 provides a summary of the evidence for traditional and structured interview validity after correction for unreliability of the criterion.

Table 2.1 Corrected correlation coefficients of structured and unstructured interviews

Study	Structured	Unstructured
Latham et al. (1984)	0.87 & 0.82	
McDaniel et al (1987)	0.45	0.36
Weisner & Cronshaw (1988)	0.45	0.36
McDaniel et al (1994)	0.44	0.33
Huffcutt & Arthur (1994)	0.54	0.20
Conway et al (1995)	0.67	0.34
Cortina et al (2000)	0.44 -0.63	0.20 - 0.35

Potential sources for bias in addition to errors of recall and subjective intra-rater and inter-rater interpretation of responses include the possible effects of attraction or similarity between the candidate and interviewer and the possibility of adverse impact from interpersonal prejudice or stereotypes. Arvey (1979), and Reilly and Chao (1982) present evidence that women with the same credentials as men tend to suffer adverse impact in unstructured interviews. This effect was moderated by the type of job they were applying for but this may be expected to relate to stereotypes of gender-related/differentiated occupations and management would be very possibly be included in this. Kinicki and Lockwood (1985) report that the suitability of an interviewee was significantly predicted

by three factors in their study, applicant attraction to the interviewer, same/different gender, and interview impression or perceived ability to express ideas, job knowledge, appearance and drive.

Following on from Hovland and Wonderlic (1939) it has become widely accepted that the validity of interviews can be greatly improved if questions are standardized. Structured interviews involve the standardization of judgements and questions and the introduction of rating scales and checklists. This procedure may benefit from job analysis in order to develop relevant and valid items and to limit the collection of irrelevant information. Examples of structured interviews include the Patterned Interview (McMurray, 1947), the Situational Interview of Latham, Saari, Pursell, and Campion (1980), (see table 1.1.2 for validity coefficients) The Patterned Behaviour Description Interview of Janz (1982); and the Structured Behavioural Interview (Motowidlo, Carter, Dunnette, Tippens, Werner, Burnett, and Vaughan, 1992). Such procedures usually report high validities but often use small samples.

Cortina, et al. (2000) explore the issue of interview structure further. Using Huffcutt and Arthur's (1994) distinction of three levels of structure Cortina et al employ hierarchical regression to uncover the relationships between interview score, conscientiousness, cognitive ability and supervisor ratings of performance. Interview structure was organised as follows: no standardisation of questions and scoring; low to moderate constraints on questions and / or scoring; and moderate – large constraints on questions and scoring. Using this typology Cortina et al report incremental validity of structured interviews over measures of cognitive ability and conscientiousness. Furthermore, the incremental validity of the interview improved as the level of structure increased. These results

suggest that the selection interview captures skills or attributes not collected through traditional psychometric measures of ability and conscientiousness. Cortina et al support the views of Dipboye & Gaugler (1993) that increased structure in the selection interview results in a valid and legally defensible technique that is less prone to individual differences bias and error associated with recall and social attribution processes.

Barclay (2001) asserts that behaviourally-oriented structured interviews provide enhanced skills training for interviewers, improved reflections upon performance, better feedback and opportunity for skill development among candidates, and offer the candidate greater opportunity to explain and discuss individual skills and working history than a traditional unstructured interview. Despite the issue surrounding the cost of training interviewers, and the need for s standardised and reliable scoring system, Barclay concludes that the behaviourally structured interview offers good reliability and criterion-related validity as well as impressive levels of face validity and satisfaction among candidates and employers.

Overall, it seems that increasing the structure of the interview improves validity and reliability. "Standardized questions provide a more consistent sample of applicant performance" (Conway, Goodman and Jako, 1995:575) and/or "provide a more job-relevant sample of applicant performance... interviewers can evaluate the information more successfully" (Conway et al.1995: 575).

Biographical data presents all candidates with a structured and standardised set of questions at the pre-selection stage. The inclusion of biodata items that have demonstrated significant prediction of the performance criterion, is likely to produce validity coefficients

in excess of the unstructured interview by the same token that increased structure in the interview improves predictive performance. Biodata may offer incremental validity over the unstructured and structured interview, as it is likely that biographical life history items may capture additional variance not typically captured by interview questions. Schmidt and Hunter (1998) present preliminary evidence from a series of meta-analyses in support of this argument.

2.2.2 Assessment centres

The term 'assessment centre' usually refers to the use of several selection methods with groups of candidates over the course of three to five days. Finkle (1976) describes assessment centres as group-oriented activities providing a basis for judgements or predictions of behaviours believed to be relevant to job performance in an organisational setting. Finkle outlines four defining characteristics of assessment centres including the assessment of groups (usually of around twelve candidates); assessment by groups; the use of multiple measurement techniques with strong emphasis on situational exercises; and the special appeal of this method for managerial selection. Testing techniques often include activities such as interviews, psychometric and projective testing, situational exercises, peer and self-ratings.

Previous research suggests that assessment centres are used by approximately 20 per cent to 37 per cent of UK organisations for managerial selection (Bridges, 1984; Robertson & Makin, 1986; Shackleton & Newell, 1991). Shackleton and Newell (1991) report that 60% of British companies used assessment centre-type exercises in selection and assessment across occupational level and function.

Cooper and Robertson (1995) argue that the use of multiple methods of assessment and multiple assessors provides a reasonably objective all-round view of each candidate (an "overall assessment rating"), which is considered to provide a good prediction of the future performance of an individual in those dimensions assessed. The main use of assessment centres is for the identification of managerial potential and the use of simulated tasks to achieve this appears to offer ratings of performance of directly job-related behaviours.

Although known to be expensive and time consuming, assessment centres may be popular due to the apparent face validity of situational exercises and the perceived increase in the chances of obtaining reliable scores and ratings from several assessors across a number of tasks. Evidence in support of the predictive validity of assessment centres is vast. For example, Handyside and Duncan (1954) report that an assessment centre used for identification of candidates for promotion within a Scottish engineering organisation proved to be a valid indicator of later promotions and job performance ratings. This study used a small sample. Further confirmatory evidence comes from Hinrichs (1978) eight year follow-up presenting results that overall assessment centre ratings correlated significantly with management level attained equally for male and female participants. Bray and Grant (1966) report data from a sample of 422 candidates taking part in an assessment centre over three and a half days. Candidates were required to complete an in-tray exercise, an interview, a problem simulation, a group discussion, projective tests, and psychometric questionnaires. The results demonstrated good predictive validity and follow-up studies revealed that 64% of those predicted to reach middle management had done so when managerial level eight years after initial assessment was used as the criterion

(Bray, Campbell, and Grant, 1974). Bray (1982) later compared assessment ratings with promotion criteria (AT&T study) and reported that 58% of those predicted to reach a third level of management status did so.

Table 2.2 The predictive validity of assessment centres

Study	Rater	Predictive coefficient
Vernon (1956)	Observer	0.44
	Psychologist	0.49
	Chairman	0.49
Cohen et al. (1974)	Overall Ratings	0.33
Cascio and Ramos (1984)	Overall Ratings	0.37
Hunter and Hunter (1984)	Overall Ratings	0.43
Schmitt et al. (1984)	Overall Ratings	0.43
Gaugler et al (1987)	Overall Ratings	0.37
Feltham (1988)	Overall Ratings	0.37
Tziner, Meir, Dahan, & Birati (1994)	Supervisor	0.24, 0.18, 0.26, 0.18

Despite the evidence in support of the predictive validity, several studies have questioned the construct validity and the possibility that assessment centres offer clinical or subjective ratings of candidates based upon opinion or judgement alone. Wollowick and McNamara (1969) point out that assessors may resist instructions to base judgements upon task-specific performance, preferring to use personal impressions and judgements in overall ratings of candidates, reducing the predictive power of the procedure. Large variations have been uncovered in the correlation coefficients of assessment centres assessing different groups of participants (Schmitt, Schneider, and Cohen, 1990).

Klimoski and Strickland (1977) question the extent to which assessment centres actually predict job performance. Many studies following this one have reported better prediction of potential in comparison to job performance, (for example, Gaugler, Rosenthal, Thornton & Bentson, 1987; Feltham, 1988; 1989). Klimoski and Strickland assert that assessment centres suffer from restriction of range, and possibly criterion contamination

because they assess suitability for promotion or selection. Candidates will naturally have been invited to attend on the basis of their suitability. Klimoski and Strickland (1987) suggest that assessment centres predict promotion that is largely dependent upon assessment centre performance and ratings in predictive studies, representing extreme criterion contamination.

Much debate has centred on the common practice of using supervisor ratings as the criterion for the validation of assessment centre ratings. Gaugler et al (1987) conclude that this practice results in flawed validity evidence and the possibility that the assessors rate participants on the basis of having the characteristics of a "good manager" rather than actual performance on dimensions. This theory explains the differential validity of ratings of potential and performance. Hesketh and Robertson (1993) conclude that assessors should be more explicit about the theoretical basis of their work and pay more consideration to the impact of assessment centres. Criterion issues include the effects of performance of differential exposure to training upon performance. Tziner, Meir, Dahan, and Birati (1994) assert that criteria should incorporate "hard data such as monetary bonuses for increased productivity, and standardized performance ratings formulated by supervisors with no knowledge whatsoever of their subordinate's performance in the assessment centre" Tziner et al.1994: 229).

Cook (1993, 1998) asserts that conclusions of assessment centre validity have been based upon a small number of studies and may therefore suffer from restriction of range. Further replication of studies and cross-validation are required. Potential moderators of validity include criterion contamination; ipsativity among participants, where dominant participants receive more favourable ratings than less dominant participants and this is not

related to actual job performance or potential; and possible influence from bias and leniency in ratings. One potential problem with assessment centres is that they may advantage participants with a strong tendency for assertiveness or surgency. Less extroverted candidates may be disadvantaged as a result. Crawley, Pinder and Herriot (1990) report that correlations between personality attributes and assessment decisions were generally low and suggested that some attributes permitted competencies to be expressed in behavioural requirements of assessment centres. For example, seven of nine dimensions tested for at the assessment centre were found to relate to the assertiveness scale of the Occupational Personality Questionnaire (Saville and Holdsworth, 1985). The implications for assessment centres involves the possible change of focus to tests of cognitive and intellectual abilities in cases where candidates show less evidence of personality attributes that permit the expression of desirable assessment centre behaviours.

The issues of convergent and discriminant validity also require further consideration, as the evidence appears to suggest that candidates should be rated by exercises rather than dimensions (ratings for dimensions assessed during the same exercise have been shown to be consistently similar).

Investigations into adverse impact or differential validity of assessment centres for female employees, ethnic minorities, and older participants have generally produced positive results against criticism of adverse impact. Huck and Bray (1976) report some adverse impact of the assessment centre procedures on ethnic minority participants, although job performance was predicted equally well for ethnic minorities and majority group members. Schmitt and Hill (1977) did not find significant differential validity based on sex or/and race. Ritchie and Moses (1983) conducted a follow up of 1600 female entry

level managers in Bell Telephone Company and reported that ratings of mid-management potential predicted rank 7 years later for females as well as for men at AT&T. Walsh, Weinberg and Fairfield (1987) found that among 1035 applicants for financial services sales posts, women received better assessments from male only assessor panels. Gaugler, et al (1987) conclude that assessment centre validity is enhanced when the percentage of females attending an assessment centre is relatively high, when the assessors were primarily psychologists, when the proportion of minority group members was low, when several assessment devices were used, and when peer assessments were included in the tasks (Gaugler et al., 1987:493). Huck and Russell (1989), and Russell and Domm (1990) also present supportive evidence of the predictive or criterion-related validity of assessment centre evaluations of managerial skill.

The transparency of assessment dimensions and criteria to certain candidates and not others may compromise the reliability and validity of assessment centre ratings. Kleinman (1993) concludes that evidence of inter- and intra-individual differences in recognizing what was being assessed has implications for the construct and criterion validity of assessment centres. This could result in discrimination against those failing to recognise the criteria for assessment.

Robertson and Iles (1988) outline doubts about the utility of assessment centres. Specifically, they assert that the assessment centre could be too expensive for wider applications in smaller organisations. Robertson and Iles also argue that important managerial roles and behaviours may not be attended to in the current format and design of many centres, especially in senior management selection.

Assessment centre ratings may be difficult to interpret. Multiple sources of rating contamination from the criterion and the rating process itself mean that it may be doubtful that the "true" relationship between assessment centre valuations and job performance can be established (Turnage & Muchinsky, 1984). This criticism is supported by the research of Klimoski and Brickner (1987). In this study, criterion contamination, self-fulfilling prophecy and the absence of an adequate definition of managerial intelligence cast doubts upon the apparent validity of assessment centre predictions of performance. Smith (1991) also picks up on this point and concludes that disappointing validity coefficients emerging from meta-analyses are making many of the users of assessment centres question the value obtained from such a costly exercise.

Attempts to improve the validity of the assessment centre follow similar patterns to improving the selection interview. Minimizing the complexity of the assessor's task may serve to increase the validity of assessment centres. Reilly, Henry, and Smither (1990) found that the application of a behavioural checklist for candidate ratings effectively reduces the difficulty that the assessor may have in attempting to attend to various dimensions simultaneously and improves the convergent validity of the overall assessment rating. This procedure is likely to work in the same way that a standardised set of questions during structured interviews makes the cognitive focus of the interviewer clearer and may reduce tangential drifts in attention. Using the same reasoning, biodata is likely to show incremental validity over assessment centres as a longer-term focus on life history events and achievements may produce a better prediction of future performance.

Feltham (1988) suggests that the statistical combination of all ratings obtained during the assessment centre would increase the predictive validity of the overall rating. The

superiority of mechanical/statistical ratings over clinical or judgmental methods has been fairly well established previously (e.g. Sawyer, 1966). Applying this principle to the assessment centre context would reduce some of the variability that inevitably forms part of subjective and clinical interpretations and ratings. Similarly, the development of a biodata predictor scale to capture the range of aspects that are assessed during an assessment centre is likely to reduce the error associated with variable subjective ratings and lend support to the prediction that biodata may demonstrate incremental validity over assessment centres.

Work sample tests are frequently used in assessment centre designs to assess individual performance on job-specific tasks. Work sample tests commonly employed in managerial selection include in-basket tests, presentation exercises, group discussions, and decision-making exercises used in assessment centres. Candidates are required to respond to items within a restricted time period and are scored on decision-making, business awareness, and other dimensions. Hunter and Hunter (1984) present acceptable criterion related validity for work sample tests and argue that the test has high face validity for both candidates and assessors. Disadvantages of work sample tests include the necessity to update the tests each time a component of the real job functions changes, and the assumption that the tests measure aspects of character, ability or personal motivation that could not be assessed more cost effectively in other tests (Cooper and Robertson, 1995; Robertson and Iles, 1988). Robertson and Iles (1988) asserted that more attention should be focused on developing specific managerial work sample tests because they function as a realistic job preview as well as a selection tool and aid self-selection. The work sample represents an example of the "sample" approach to obtaining predictive measures of performance (Wernimont and Campbell, 1968) and may be used for selection and training purposes.

Consistent evidence of good validity and reliability of work sample tests in selection has been attained throughout the available research (Dunnette, 1972; Robertson and Kandola, 1982; Schmitt, Gooding, Noe, and Kirsch, 1984). Advantages of the work sample test include the increased face validity, and plausibility to applicants. The use of work sample may appear to be limited in application to people with existing job skills although trainability tests can assess how well applicants may learn new skills (this approach is widely used in the UK).

The leaderless group discussion is another frequent component of managerial assessment centres. In this task candidates are requested to discuss a topic (job-related for better validity and reliability and to attempt to exclude potential sources of bias) for a restricted period of time. Uninvolved observers rate the contributions of each candidate. For example, in an IBM assessment centre, candidates were required to make a five-minute oral presentation of a promotion decision and then defend their decisions in a leaderless group discussion. Observers allocated ratings on dimensions of aggressiveness; persuasiveness, oral communication skill; self-confidence; resistance to stress ("effectiveness"), energy level; and interpersonal contact (Wollowick & McNamara, 1969). Cooper and Robertson (1995) report that leaderless group discussions assessed managerial ability on the basis of applicants' ability to get on with others, influence others, think clearly and logically, and apply experience to new problems. Overall, these dimensions were translated into:

- Interpersonal skills,
- Leadership potential

- Business awareness
- Problem solving

Disadvantages of the leaderless group discussion test include the possibility of social influence impacts upon assessor ratings of candidates, the possible introduction of bias or discrimination into the topic, and the influence of personality variables over group roles and assertiveness in the group situation (see Crawley, Pinder, & Herriot, 1990).

In summary, the assessment centre is likely to offer the most comprehensive assessment of candidate suitability. Where this method incorporates a structured interview, a work sample test, and peer nominations, it represents the selection method with the most potential of revealing impressive predictive and face validity. The costs of assessment centres when considered in terms of a utility model are compensated directly from the benefits of selecting able and enthusiastic managerial staff (for example, Cascio and Sibley, 1979; Hunter and Hunter, 1984). Tziner, Meir, Dahan and Birati (1994) report that even assessment centres with only modest predictive validity can prove to have utility and may "outperform traditional methods in detecting high-level management potential" (Tziner et al.1994: 240).

2.2.3 Cognitive ability measures

A cognitive or mental ability test is any form of standardised psychometric test measuring or assessing intellectual functioning. Such tests may provide a measure of general intelligence and may be general or specific to job tasks or requirements.

Approximately 5% of the British population of applicants could expect to complete psychometric tests in selection through the 1980's (Smith, Gregg, and Andrews, 1989). The majority of these applicants are graduates and may be applying for managerial work. The use of mental ability measures in selection involves the decision to use test batteries or individual test, and to incorporate measures of specific or general abilities the use of different tests for different occupations has recently become more popular. For example, Pearlman, Schmidt, and Hunter (1980) tested secretaries across a number of organisations. Validity generalisation analysis uncovered no residual variance between jobs in different organisations indicating that similar tests may be valid for a majority of clerical occupations without the need for a local validation exercise or extensive job analysis methodologies.

As early as 1928, Hull argued that cognitive ability tests measuring specific abilities offer greater predictive validity of job performance than general ability or aptitude tests. This has contributed to the tradition of using multiple test batteries in American selection. However, Ghiselli (1966, 1973) asserts that managerial performance can be best predicted by general intelligence tests and perceptual ability tests (r=0.53 general intelligence; r=0.43 perceptual ability). Hunter and Hunter (1984) support that pioneering work of Ghiselli (1966, 1973) and conclude that tests of mental ability offer the highest predictive validity of job performance.

Korman (1968) reported that measures of verbal ability offer fair predictions of potential for first line supervisors but not for higher-level managerial performance indicating the need for consideration to the level of work. In contrast, Grimsley and Jarrett (1973, 1975) reported that mental ability test scores and self-description inventories could reliably

distinguish between top and middle level managers. Grimsley and Jarrett argue that differences between these levels of employment reflect differences in mental ability and personality rather than the influence of job-experience or knowledge.

One key advantage of measures of cognitive ability is validity generalisation. This refers to the ability of a test to demonstrate predictive validity across organisational and specific job role boundaries. Hunter & Hunter (1984) report that mental ability tests consistently demonstrate validity generalisation, one outcome of this research is the presumption that well-developed tests will show at least some correlation with job performance in almost all jobs. Schmidt, Hunter, and Raju (1988) support this generalisability of validity. Murphy (1994) similarly reviewed a large number of meta-analyses of the validity of cognitive ability tests as predictors of performance with favourable results. However, job complexity may moderate the validity of tests across different occupations. It appears that ability tests show higher validities in more cognitively demanding jobs (Hunter & Hunter, 1984).

Assessment of the usefulness of various tests requires evidence of content, criterion-related validity, discriminant and convergent validity, incremental validity, and construct validity. Factors affecting test reliability include test length, objective marking, clear test instructions, standardized testing conditions, etc. Test norms are needed to give psychological meaning to subjects' scores. Test interpretation in comparison to such norms can only be as good as the norm samples are representative and adequate sample sizes have been used for the collection of standardization data.

Dunnette (1972, cited in Cook, 1993) investigated the validity of mental ability tests for various occupations within the American Petroleum industry. Most of the correlations were between 0.20 and 0.30. Some tests showed greater validity than others.

Hunter and Hunter (1984) present data showing that cognitive ability tests predict supervisory ratings and training success in a variety of jobs well, and predict content-valid work-sample performance even better. Most of the variance across validity studies of mental ability tests was attributed to sampling error. Cognitive ability tests are said to predict performance for all job families.

Criticisms of the available validity data have been levelled at the assumption of Hunter and Hunter (1984) that supervisor ratings (used as the criterion) have an average validity coefficient of 0.60. Hartigan and Wigdor (1989) similarly estimate the supervisor rating validity at 0.80. It should also be noted that validity generalisation analyses estimate unreliability and restriction of range and cannot directly take account for other potential moderators. It is worth bearing in mind the "File Drawer phenomenon", a tendency for insignificant findings to remain unpublished. With this in mind, meta-analyses based upon published research only may represent a somewhat biased view of all of the actual research findings. Schmidt (1992) asserts that approximately 80-90% variability in test validities is due to statistical artefact and points out that large-scale meta-analytic reviews of the validity data cannot improve upon the base of research articles they are based on.

The trend for using mental ability tests decreased throughout the 1960's partly due to controversial debates surrounding research and conjecture into the heritability of mental ability (e.g. Jensen, 1969). Increased concern about equal opportunities throughout the

1960s and seventies also led to a decline in popularity and an increase in the notion of "competencies" for work. Hunter and Hunter (1984) conclude that tests providing valid predictions of performance for one group may provide valid predictions of the job performance of other racial groups in the same way. Evidence suggests that the test score differences between racial groups are likely to reflect real differences in educational achievement and actual job performance. Selection techniques would therefore be developed along lines of proportional/almost proportional representation of minority groups. Whether this operates at all levels of occupational status and across organisations, and explanations relating to the aetiology of intergroup differences in performance are beyond the scope of this discussion. The available evidence appears to confirm that a hypothesis of "differential validity among racial and gender groups should be rejected" (Cook, 1993).

2.2.4 Personality measures

Robertson and Kinder (1991) support the findings of Campbell, Dunnette, Lawler, and Weick (1970), and identify cognitive and non-cognitive factors as important determinants of managerial success. Despite the low /moderate low correlation coefficients obtained for personality constructs as single predictors of job performance (for example, Ghiselli, 1973; Reilly & Chao, 1982; Hunter and Hunter, 1984; Johnson & Blinkhorn, 1994), personality items may predict variance not accounted for by predictors of cognitive ability (Gellatly, Paunonen, Mayer, Jackson, & Goffin, 1991; Robertson & Kinder, 1993). The research presented in this thesis will contribute to the existing literature on personality and managerial job performance and interest and extends the view of the manager as an 'independent, stable extrovert' (Bartram, 1992: 168).

During the 1980s the Occupational Personality Questionnaire was developed and standardized by Saville and Holdsworth (Saville & Holdsworth, 1985). This test is specifically designed for industrial settings. Robertson and Kinder (1993) assessed the criterion related validity of the OPQ and found correlations from 0.09 to 0.32. The authors of a variety of personality inventories have developed methods of applying scales to specific occupations. For example, the 16PF includes equations for the calculation of leadership potential, ability to adapt to a new job, and the type of team role that a person is best suited to or prefers (Smith, Gregg, & Andrews, 1989).

Ghiselli and Brown (1955) suggest that there is a low relationship between vocational interest and job performance, this view has been supported by Kline (1992) who argues that interest is not a good predictor of job success and should not form part of the selection process. However, a variety of vocational interest scales have revealed interesting results. For example, Mahoney, Jerdee, and Nash (1960) compared effective and ineffective managers on the California Personality Inventory, The Strong Interest Inventory, The Wonderlic Personnel Test and a biodata questionnaire. They concluded that effective managers could be discriminated and are more intelligent than ineffective managers. S/he tends to be aggressive, self-reliant, has more educational training, more active in sports during youth. Similarly, Harrell (1972) studied Stanford MBA students five years after graduation and reported that the Guilford Zimmerman Temperament Survey (GZTS) Ascendance dimension predicted salary level positively in large firms (more than 1000 employees). The 'Social Interest' and 'General Energy' scales also predicted success.

Bendig (1963) investigated the relationships between extroversion-introversion and neuroticism-adjustment as measured by the Maudsley Personality Inventory, with

vocational interest. He reported that the Business scale of the interests inventory (from the Strong Vocational Interest Blank) was negatively correlated with neuroticism for male and female college students, and negatively correlated with extroversion for male students only. Further results from Bendig (1963) are summarised in Table 2.3.

Table 2.3 Summary of Bendig (1963)

	Men	Women
Extroversion		
(Positive correlations)	Public Administration	Personnel Director
	Sales Manager	Office Manager
	Life Insurance Sales	Sales Manager
Neuroticism		
(Negative correlations)	Personnel Director	Personnel Director
	Office Manager	Office Manager
	Sales	Sales
	Public Administration	

Bendig concludes that male students with vocational interests in management tended to be stable introverts. Both extroverted and introverted female participants exhibiting relative stability also appeared to have business/managerial interests. It is important to note that this research was conducted in the 1960s, which may at least partially explain the different results found in recent explorations of managerial personality. The nature of managerial work is likely to have changed considerably over the last thirty years.

The decrease in traditional manufacturing industry and the increase in telecommunications and media may have resulted in the apparent need for more extroverted managers. Barrick

and Mount (1991) report a positive relationship between extroversion and managerial performance. Along similar lines, Rim (1961) reported a moderately strong correlation between extroversion and high need for achievement and low fear of failure. It might reasonably be said that such characteristics would serve managers well in competitive business.

Day and Silverman (1989) investigated the relationships between cognitive ability, orientation towards work, degree of ascendancy, and degree and quality of interpersonal orientation among accountants. Results suggested that these personality dimensions show promise for predicting the job performance of accountants. This led Day and Silverman to suggest that job-relevant personality measures be used in conjunction with cognitive measures in selection.

In support of Campbell, Dunnette, Lawler, and Weick (1970), Robertson and Kinder (1991) identified both cognitive and non-cognitive factors as important determinants of managerial success. Robertson and Kinder (1993) conclude that the use of personality tests in selection may provide information related to success that is not provided by traditional cognitive tests. The data from Robertson and Kinder have since been reanalysed by Salgado (1996) who points out that they had miscalculated the sample error, underestimating the relationships of the criterion-related validity of personality selection.

The findings of Gellatly, Paunonen, Meyer, Jackson, and Goffin (1991) offer further support for the use of personality inventories. Gellatly et al evaluated the contributions of cognitive ability, personality and vocational interest to job performance for first line

managers. They reported that personality and interest measures provided information not given by the cognitive ability measures.

Jackson and Rothstein (1993) criticise the current trends for using personality measures in selection and validation research. Their objections can be summarised as follows; a small number of traits have been studied and been followed by huge generalisations; validity analyses and conclusions have been reached after the analysis of inadequate samples of the available research literature; statistical anomalies and error seem to characterise this field of research. Jackson and Rothstein propose that providing well constructed and validated performance measures are employed as criteria (guided by job analysis and predictive hypotheses), along with appropriate statistical analyses, personality measures may offer utility for personnel selection. They argue that the negative results achieved by Blinkhorn and Johnson (1990) in their reviews of the 16PF and the CPI are unsurprising because both inventories are grounded in "antiquated technology and theory". Zuckerman (1989) had previously warned against the low internal reliability and criterion validity of the 16PF. Eysenck (1989) has criticised the CPI because of the absence of information about item intercorrelations in the test manual. Jackson (1990) lists the preconditions for obtaining valid predictors of job performance from personality measures. The measure/s must demonstrate independent evidence of construct validity and freedom from bias, and attention should be directed at overcoming or accounting for the unreliability of the criterion.

Dunn, Mount, Barrick, and Ones (1995) report that general mental ability and conscientiousness were the most important attributes to managers making selection decisions. The sample size in Dunn et al.'s (1995) study is small and conclusions based

upon research using hypothetical job applicants cannot be considered to be widely generalisable.

2.2.5 Biographical data

Although there is no accepted consensus on the nature of information collected in biographical life history forms, biodata items typically include age, sex, marital status, educational achievements, job history, hobbies, and other items typically found on application forms (Cook 1993). Biodata questionnaires collect self-reported information regarding the demographic, academic achievement, work experience, interests and personality attributes of individuals (Owens, 1976).

Mael (1991) proposes a taxonomy that emphasizes the importance of verifiable information on controllable life experiences. McCormick & Ilgen (1980) observe two types of biographical information, 'job-related' and 'labour market related' items. Drakeley, Herriot and Jones (1988) assert that biodata should involve a combination of achievement, background and commitment items. Drakeley et al point out that although these categories are not discrete, each category relates differentially to life experience and opportunity and considers the role of socio-economic factors on background information.

Asher (1972) distinguishes "Hard" items, those requiring factual and verifiable information and "soft" items, which may be easier to fake but are less intrusive and more acceptable to candidates (Cook, 1993). "Hard" biodata items include questions on educational qualifications, age and marital status. Soft items may ask for individual judgments or attitudes and opinions and are not directly verifiable.

A factor analysis study by Baehr and Williams (1967) presents five major categories of biodata items; educational background and achievement, upward mobility and drive, personal-social leadership, financial achievement and background, stability; and status quo orientation.

It is worthwhile discussing briefly issues surrounding the apparent lack of popularity of this method, which provides impressive validity figures (refer to table 2.4). Biodata has traditionally been developed for use within large organisations recruiting vast numbers of applicants annually. In Britain, the civil service represents such an institution. Biodata inventories were developed here for use as a screening device in the early stages of selection. Inventories consisted of items that are found to statistically predict performance on specific dimensions of (managerial) work.

Robertson and Iles (1988) present a review of the validity evidence from recent research and discuss the evidence that biodata inventories could successfully predict performance and voluntary withdrawal from training (Drakeley, Herriot and Jones, 1988). Possible explanations for the limited use of such an apparently valid instrument include:

1. The considerable expense involved in the development of organisation-specific biodata inventories,
2. The apparent need for large samples for the validation of questionnaires,
3. Potential lack of face validity and lack of acceptance by candidates,
4. The possibility that traditional inventories may contain items reflecting access to status hierarchies and opportunity structures (leading to discrimination and bias),
5. The fact that biodata appears to present a "static" or "fixed" view of an individual predetermined by previous events,
6. The possible time-limited efficacy of items.

Although the costs of developing and validating a biodata inventory are relatively high, organisations with a large turnover of staff may find that biodata has considerable utility in the first stages of the selection process in comparison to time-consuming techniques such as assessment centre exercises and interviews. Robertson and Makin (1986, 1993) report that approximately 6% of British organisations use biodata. The largest organisations in the UK appear to use biodata most often, for example, the UK Civil Service Commission have a biodata inventory for tax inspectors, British Airways for cabin crew selection.

Ryan, Baron and Page (1999) report popularity of selection methods from a sample of 959 organisations across 20 countries. Using a response scale rating from '1' = 'never use', to '5' = 'almost always use', the relative low use of biodata reveals the unpopularity of biodata for UK selection.

Table 2.4 Popularity data for UK selection methods (Ryan et al., 1999)

UK data (108 organisations)	Mean	S. Dev.
Biodata	1.23	0.71

Interviews	3.88	1.53
References	4.37	1.01
Personality & work styles inventories	3.46	1.37
Cognitive ability tests	3.08	1.52

Among the countries included in this research Greece had the most frequent use of biodata (n = 27, mean = 3.87, standard deviation = 1.63, followed by Germany (n = 35, mean = 2.77, standard deviation = 1.72). These results are discussed with reference to cultural 'avoidance of uncertainty'. Of the selection methods discussed, biodata had the strongest positive correlation with avoidance of uncertainty, demonstrating that cultures and organisations keen to learn as much as possible about candidates are perhaps more likely to utilise biodata in addition to other methods of selection. Organisations from Greece received an 'avoidance of uncertainty' index score of 112 in comparison with a score of 35 for UK organisations (mean 'Avoidance of Uncertainty' score was 54.63, Ryan et al., 1993:365). Larger organisations were more likely to report using a variety of approaches. A follow-up study would be useful in investigating repeated use of each of the selection methods. This research could assess organisational and cultural views on the efficacy of each approach to avoiding uncertainty, thereby exploring potential growth in frequency of use and predicting larger cultural shifts in employee selection.

In summary, biodata has been successfully used to predict performance on job-specific criteria (e.g. Ghiselli, 1966, 1973; Asher, 1972; Owens, 1976), turnover, adjustment, accident proneness, job satisfaction, team performance (Stokes & Cooper, 1994), organisational identification (Mael & Ashforth, 1995), and creativity (Barge & Hough, 1988; Childs & Klimoski, 1986; Drakeley et. al., 1988; Rothstein, Schmidt, Erwin, Owens, & Sparks, 1990; Smith, Albright, Glennon, & Owens, 1961).

2.3 Summary of approaches to managerial selection

As Gersein and Reisman (1983) conclude, selection decisions should be guided primarily, by an assessment of risks and rewards of the procedures used and of the ability and familiarity of the selection personnel in carrying out these methods. The implications of the available research point to the introduction of strategic selection techniques founded on a starting point that incorporates job demands in the context of generalised strategic requirements. A precise and straightforward method for documenting job requirements can then provide a vehicle for communication and a common language for the discussion of job functions and tasks and individual abilities required for the performance of these. Gersein and Reisman (1983) provide a model for strategic selection based upon a job-requirements approach. Through the processes of job description and person specification, (e.g. Wilkinson and van Zwanenberg, 1994) methods of selection can themselves be selected in accordance with the purposes of criterion measurement and the prediction of behaviour.

Gulliford (1991) asserts that the realistic assessment of managerial potential requires the objective measure of individual aptitude and/or characteristics such as personality and cognitive ability and the understanding of the opportunities in which it is likely to be expressed. Similarly, Baehr and Orban (1989) used personality and mental ability tests to predict salary in 800 line managers, salespersons, professionals, and technical experts and conclude that personality results (used in combination with cognitive ability measures) demonstrate incremental validity over mental ability test results for managerial workers.

If the nature of effective managerial work is amenable to investigation and distinction, there are several potential uses for this information. Performance appraisal, career counselling and training and managerial selection all stand to gain from the delineation of a 'prototype' or profile of the effective manager. However, such a profile could be applicable across organisations and industries and must be based on a person specification, in turn the outcome of the delineation of an 'effective manager.'

The use of psychometrics and personality measures in particular does imply acceptance of a trait conceptualisation of personality. The cross-situational consistency debate concerns the extent to which individuals exhibit stable patterns of behaviour over time and different situations. Mischel (1968) has criticised the utility of mental tests and asserted that tests of personality and to a lesser extent, mental ability tests measure internal traits or dispositions that are actually inconsistent and may not exist. However, Kenny and Zaccaro (1983) analysed data on leadership using recombined groups of people to observe how patterns of leadership emerge. Social psychology has long assumed that leadership was not a trait per se but a role that individuals may take on in certain circumstances. Kenny and Zaccaro reported that individual identity accounted for between nine and 82% of variance in rotation studies. This suggests that leadership may be a trait as well as a situation role and also that some individuals demonstrate a personalised trait or adapt a specific situation more than others. In short, the results show unequivocal ambiguity towards any clearer understanding of the root or origins of management in terms of a certain person or certain situation. The researcher is faced with versions of the chicken and egg saga as well as the 'what' versus 'who' one knows conundrum. Inevitably the questions lead back to the issues surrounding entry into management, in pursuit of the

trainee. To this end, biographical life history information may be enlightening in illuminating formative life events beyond the remit of traditional measures of personality and reasoning.

The use of biodata, which could flow from the combined collection of life history information, cognitive and personality data, offers the best indication of likely job performance. The superior predictive validity of mental ability tests (e.g. Baydoun & Newman, 1992) and the extra information provided by personality measures can highlight the individual differences likely to moderate performance and job satisfaction.

Ideally, managerial selection would combine a content-valid, criterion based mental ability test (Rose and Baydoun, 1995) with a content-valid personality measure that controls for/takes account of the "ideal employee" frame of reference. The benefits of utilising such a strategy include the development of a valid and legally defensible method of selection at relatively low costs (in comparison to alternative methods of managerial selection such as assessment centres).

Chapter 3 Biodata

3.1 Approaches to biodata development

3.1.1 The empirical approach

Furnham (1997) outlines the history of biodata, beginning with Goldsmith (1922). Goldsmith designed a Weighted Application Blank (WAB) for the selection of insurance salesmen. The classic WAB is 'invisible' to the applicant (expecting a standard application form) and was heralded as an unfakeable selection method. More recently, biodata has been collected using self-report multiple-choice questionnaires and the issue of response distortion remains contentious and will be discussed in section 3.2.

The empirical or traditional method of biodata development involves the attribution of weights to questionnaire items on the basis of relative contribution (or predictive validity) to a criterion under investigation (Mael, 1994). This method follows on from the Weighted Application Blank tradition. The WAB design was developed on the basis of an "empirical association of the response options to some criterion in order to provide the best prediction of that criterion"(Childs & Klimoski, 1986:4). Hogan (1994) points out that the purely empirical methods of biodata development are most commonly operationalised where the prediction of an external criterion is the primary goal. Hogan discusses the importance of cross-validation for this type of biodata. A major problem associated with empirical keys is the potential lack of generalisability across organisations and over time.

Empirically developed biodata have been criticised for an apparent lack of generalisability of biodata keys beyond the organisation in which the key was developed, and beyond the criterion against which it was validated (e.g. Hunter & Hunter, 1984). Research conducted by Laurent (1970) and by Rothstein, Schmidt, Erwin, Owens, and Sparks (1990), suggests promise for generalisable biodata keys although it should be noted that large cross-organisational and international divisions of multi-national organisations were used in the validation and key development stages of the research.

The empirical method of employing items purely on the basis of criterion-related value has been criticised as 'dustbowl' or 'shotgun' empiricism by writers such as Guilford (1959). The term 'shot-gun' is applied because of the ambiguity surrounding the predictor-criterion relationships. Guilford (1959) asserts that empirically developed biodata, characterised by shotgun empiricism may be largely devoid of theory and generality. In addition, a purely empirical approach to item development does not advance or explain the associations between item content and job performance, and therefore extend the available theories of job performance. It has also been asserted that this method of item development capitalises on chance associations between the criterion and the developmental sample. Due to sources of variability from the criterion measure and scores, this method may be prone to error (Van Zwanenberg & Wilkinson, 1996). Similarly, Dunnette (1962) has emphasised the need for biodata theory to go beyond simple predictions and into causal relationships and, several other authors support this view (Baehr & Williams, 1968; Pace & Schoenfeldt, 1977; Wilkinson, 1995).

For the most part, empirical biodata is developed in a single organisation and consists of items validated with an index of performance such as supervisor ratings. As such,

empirical biodata may present a 'snapshot' of organisational and labour market trends and culture at a specific time. Changes in the nature of work, the criterion measure, tools used within occupations, labour market conditions, and the economic climate may act to introduce variance in a biodata key over time and reduce the temporal stability of the instrument, decreasing the reliability and therefore, the predictive validity of the questionnaire.

3.1.2 The rational approach

Asher (1972) presents the first account of a method for a rational approach top biodata model development. The rational approach is based upon the 'attempt to quantify composites of items that measure an interpretable set of constructs' (Mitchell & Klimoski, 1982). In this model, a conceptual framework or hypothesis of underlying construct relationships is the basis or guide for item construction. In support of Asher's claims, several observations have been made regarding the longevity of rationally derived biodata models. Rational models demonstrate less shrinkage of validity than empirical methods of development (Mitchell & Klimoski, 1982; Mumford & Owens, 1987; Mumford & Stokes, 1992; Schoenfeldt, 1989 as cited in Mael, 1994).

Asher (1972) suggests that items designed on the basis of a-priori theories of predictor-criterion relationships should show enhanced validity and understanding in comparison with empirically derived items. Several attempts at applying personal construct history approaches to biodata development have been made, for example in the work of Levine & Zachert, (1951), Loevinger, Glesson and Dubois (1953), Morrison, Owens, Glennon and Albright (1962), Baehr & Williams (1967), and Matteson, Osburn and Sparks (1969),

Criticisms of the rational approach to biodata have included the need for greater psychological knowledge on the part of the biodata developer. The rational model does not necessarily explore the links between items and work performance to a substantial degree. As yet, there is no generic theory of job performance. It is possible that fuller understanding of the antecedents of successful managerial performance could enrich existing knowledge. In summary, the rational approach seeks to match item prediction trends with underlying psychological theories of associations between items and predicted job performance. Evidence suggests that the rational approach, although more intuitively appealing than the traditional approach has slightly lower validity than purely empirical inventory designs (for example, Mitchell & Klimoski, 1989).

Reiter-Palmon and Connelly (2000) compare rationally and empirically keyed biodata developed from theory-based and non-theory based pools of items in the prediction of criteria including high school and college grade point average. Reiter-Palmon and Connelly outline several arguments in support of the view that rationally developed biodata should have greater validity than empirically derived scales. Firstly, rational scales are likely to show greater generalisability due to less reliance on chance or transient relationships between variables. The validity of rational scales is less likely to show shrinkage over time, as the prediction should be as consistent as the psychological attributes underlying the items are. Reiter-Palmon and Connelly (2000) criticise the conclusions of supporters of the empirical approach such as Mitchell and Klimoski (1982) for a lack of attention to the quality of the item pools from which the scales were developed.

3.1.3 Factorial and/or internal approaches

Factor analytic approaches to biodata development combine elements of rational and empirical approaches. Here a rational interpretation of factors is attempted after empirical relationships are ascertained. The rationality is 'post hoc'. Factor analysis is used to identify the main constructs that items tap, (for example, Baehr & Williams, 1968; Mitchell and Klimoski, 1982; Childs and Klimoski, 1986). An example of this is the 'sub-grouping' approach of Owens (1976) within the Developmental–Integrative model of biodata development.

The Developmental-Integrative model subgroups individuals in terms of experiential similarities. It is hypothesized that groups of individuals with similar life experiences may behave in a similar way in the future. Owens (1968, 1971) asserts that biodata items designed to identify patterns of experience and development particular to the individual can also identify or 'subgroup' that person, resulting in enhanced predictive validity of biodata factors (because the best predictor of future behaviour is past behaviour). Sub-grouping is said to be a parsimonious way of dealing with large numbers of applicants. Essentially, this conceptual model of classification (subgroups were tested for internal homogeneity and external differentiation) represents an attempt to provide a post hoc view of the developmental experiences of each and any individual. The developmental - integrative approach incorporates the 'inputs' to the organism (e.g. parental warmth, control etc.) and the 'prior experiences' of each individual. In addition, the subgroup members demonstrated different profiles on mental ability, personality, and interest tests. This method was supported by data from Brush and Owens (1979) in the classification of oil company employees.

Owens and Schoenfeldt (1979) conducted a study on a sample of 2,000 undergraduate students. Participants were required to complete a 659-item biodata questionnaire that covered aspects of family life, school related activities, interests and attitudes derived from life experiences, sports participation, and extra-familial relationships. Factors emerging from the analysis are outlined in Table 3.1. A total of 13 factors for males and 15 factors for female participants were identified. It was concluded that the majority of the participant sample could be sub-grouped according to these factors (Owens, 1976; Owens & Schoenfeldt, 1979).

Table 3.1 Sub-grouping categories (Owens, 1976)

Owens subgroup factors
Warmth of parental relationship
Academic achievement
Social extroversion
Athletic interest
Intellectualism
Aggressiveness / independence
Socio-economic status
Parental control versus freedom
Social desirability
Positive academic attitude
Religious activity
Sibling friction

Investigations of the validity of this design have focused upon the comparison of subgroup definitions and psychological test profiles, and observations of the predictive efficacy of subgroup status in performance criteria. Eberhardt and Muchinsky (1982) report evidence of the stability of Owen's (1976) biodata questionnaire factors.

Brown (1994) outlines limitations with this 'cluster analysis' approach to categorisation and discusses the problems associated with individuals who do not fit into discrete

subgroups ('outliers' and/or 'overlaps'). More than a quarter (27%) of Owens and Schoenfeldt's (1979) sample are said to present this problem. Ultimately, the sub-grouping method of Owens and Schoenfeldt may not predict performance for around a quarter of all applicants. Brown points out that loosening the subgroup membership boundaries may reduce the predictive efficacy of the original model. In addition, the developmental-integrative approach to biodata questionnaire design was designed and validated on U.S. college samples (Brown, 1994; 224).

Ashforth and Mael (1989) suggest that social identity theory may offer a fuller account of the effects of situational and experiential influences acting upon the individual, i.e. the influence of social group categorisation of the individual. Self-categorisation theory (Turner, 1991) may also explain subgroup similarities of behaviour and life experience, in terms of perceived similarity of, and emotional significance involved in social group (or category) attachments. Biodata items tap individual self-perception and therefore, the social categorisation or "social identity" of the individual, alongside group-based subjective belief structures about the permeability of group boundaries and the fixed or dynamic nature of the social status quo with regard to intergroup relations and individual career planning. These and related issues become pertinent when discussing managerial selection with reference to the relative autonomy of managerial employees and the perceived power of the managerial job incumbent.

The 'ecology model' (Mumford & Owens, 1987; Mumford, Stokes, & Owens, 1990) may be described as an attempt to integrate models of human development and vocational choice as measured by life history interviews, life history correlates of job requirements, and/or empirical findings regarding factor loadings of biodata items. Owens' (1976)

questionnaire was administered to large sample of undergraduate students. Analysis revealed reliable factor structures for males and females that were consistent across samples and previous research, this approach therefore became known as the "ecology model". This model was devised to address the issue of stability of behavioural and experiential patterns throughout individual development. It is asserted that individuals seek situations and experiences that have previously afforded positive reinforcement and reward. Similarly, it is expected that tasks associated with negative reinforcement will be avoided. The process of life experience and life choice will result in behavioural patterns and sets of characteristics that may define the individual in relation to groups of similar and different individuals. Mumford and Owens (1987) present six potential sources of information for item development.

Table 3.2 Sources of item development (The Ecology model)

Item Sources
Human development literature
Life history interviews with job incumbents
Typical factor loadings of biodata items
Known history correlates of various job specifications
Biodata items with known predictive validities
Items generated from the investigator's psychological knowledge

Russell (1994) points out that one of the problems in rational/ecological model item development are the limitations posed by the boundaries of the biodata developer's imagination. Russell concludes that despite this limitation, rational item development can bridge the gap seen in purely empirical biodata and purely rational approaches, i.e. the gap between "theory, item content, and criterion performance measures" (Russell, 1994:32).

3.1.4 Alternative approaches to biodata development

Mael and Hirsch (1993) distinguish "rainforest-empiricism" and "quasi-rational" approaches to biodata development. Rainforest-empiricism refers to a largely empirical approach to biodata item choice and keying with the aim of using some rationality in order to counter potential capitalisation on chance and shrinkage of validity associated with purely empirical methods. This approach therefore represents an attempt to utilise empirical methods combined with an awareness of the related constructs. The quasi-rational approach involves constructing biodata analogues of each of the scales on the Assessment of Background Life Experiences (ABLE) Inventory and was developed as an attempt to reduce the shrinkage of validity that may occur in purely empirical scales. Using leadership performance as the criterion, both approaches decreased the level of socially desirable responding associated with the original psychometric instrument. Mael and Hirsch conclude that the quasi-rational approach provides greater understanding of the psychological processes underlying performance, while scales derived using the rainforest-empiricism approach demonstrate the greatest level of incremental validity over the original ABLE test.

Other military research includes the quasi-rational methods of biodata model development such as that of Kilcullen, White, Mumford and Mack (1993) and White and Kilcullen (1992). Kilcullen, White and colleagues combine subjective and objective items to capture existing temperament constructs. All items are written based upon a priori constructs. In a similar vein, Hough and Paullin (1994) present a model of construct-oriented scale construction developed for Project A (within the U.S. Army). This form of biodata inventory is developed on the assumption that constructs relate differently to separate criteria of job performance and to other item constructs. Such scales have

demonstrated convergent and discriminant validity, as well as face validity (in terms of the psychological constructs being measured), this approach has also been adopted by Mumford, Costanza, Connelly, and Johnson (1996). Mumford et al (1996) question the construct and content validity of traditional biodata scales developed within rational and empirical approaches. They report that scales consisting of items developed on the basis of psychological constructs demonstrate content and construct validity as well as criterion-related validity. Item keys developed in this manner are likely to be generalisable across organisations and stable over time as far as the psychological constructs they tap demonstrate these qualities.

Wilkinson (1995) considers the application of construct-oriented biodata to non-military samples with the development of biodata analogues of psychometric measures of critical thinking ability, extroversion and occupational interest. Wilkinson (1993, 1994, 1995) asserts that construct-oriented biodata may address several of the prominent criticisms of traditional methods of biodata development and keying. Wilkinson's alternative model for biodata development, in common with the quasi-rational approaches outlined above, focuses upon the prediction of individual difference characteristics or attributes. Wilkinson asserts that focussing upon the characteristics of the candidate rather than the job could eliminate some of the potential sources for error associated with job and criterion change over time and the associated shrinkage of validity and temporal instability previously associated with biodata. Focus on the individual also affords reduced sample sizes required for the validation of biodata keys as the emphasis changes from peculiarities of doing one job in a single organisation at a specific moment in time, to personal attributes in a dynamic, career progression framework. Because this model of key construction is person-centered rather than job-centered, it may alleviate problems associated with job and/or organisational specificity, leading to increased generalisability

(Wilkinson, 1997). In summary, the alternative approach to construction of biodata models predicts personal characteristics that can then be fitted to the job and organisation rather than vice versa.

Mumford, et al. (1996) describe the applications of worker-oriented biodata. The utility of this approach to model development may be especially beneficial where candidates 'lack prior exposure to situations similar to those found on the job. There is a need to generalise predictions across multiple criteria, and the measures are used to draw inferences about the likely causes of performance' (Mumford et al., 1996: 363).

3.2 Reliability of biodata

Investigations into the reliability of biodata have generally followed the psychometric tradition. Most frequently, test-retest reliability is taken as an indicator of stability over time (often referred to as 'the coefficient of stability'). The measures of internal reliability such as spilt half or alternate form reliability often conducted on psychometric measures of unitary constructs have not typically been applied to traditional biodata inventories.

Potential sources of unreliability in biodata responding include changes in the candidate, the nature of work, the employment climate and variance in the criterion against which the biodata inventory is validated. Van Rijn (1980) reports several sources of variance including errors in memory, carelessness, and response bias in candidates.

Controversy surrounding the use of self-report techniques has included the accuracy of recall of earlier life events (Owens, 1976). Ross and Conway (1986) point out that adults may interpret and reconstruct events in terms of present personalities / present psychological situations. A further source of variance is the problem deliberate response distortion.

3.2.1 Response distortion and biodata

Klein and Owens (1965) studied the fakeability of biodata used in predicting research creativity. They report that undergraduate students could distort their responses and 'fake good'. Similarly, Kluger, Reilly and Russell (1991) report that research volunteers pretending to be job applicants gave socially desirable responses to biodata items (these findings are supported by Gunter, Furnham & Drakeley, 1993). Gunter et al argue that motivation to produce a 'good profile', failed recall, and poor introspective awareness could contribute to inaccurate or distorted profiles of candidates.

Asher (1972) asserted that 'soft' items (requiring self-perceptions, preferences, and opinions from the applicant) may be susceptible to response distortion in terms of 'socially desirable' responding (Crowne & Marlowe, 1960). Mael (1991) has suggested that the accuracy of responses may be improved if permission is requested to verify 'hard'

information. The research evidence supports this hypothesis (for example, Trent, Atwater, & Abrahams, 1986; Shermis, Falkenberg, Appel, & Cole, 1996).

Many of the concerns aired over use of personality assessment in selection apply to the issue of self-disclosure among candidates completing biodata inventories. For example, where test items are transparent, candidate responses may be directed at an image of the ideal candidate and tailored toward this. Herzberg (1954) reports that job applicants may present 'faked' good profiles to a greater extent than samples of volunteer research participants, however, Dunnette, McCartney, Carlson, and Kirchner (1962) and Hough et al (1990) report that few applicants resort to socially desirable responding.

Hough, Eaton, Dunnette, Kamp, and McCloy (1990) recommend caution in the application of personality measures to the selection context. In particular, the use of measures to detect socially desirable response sets is advocated, warning applicants that inaccurate responses will be detected and disregarding of any inaccurate or carelessly completed personality measures. Most inventories use an endorsement format (true/false, yes/no). This approach permits rapid completion by the candidate but may encourage response sets where subjects repeatedly select a positive or negative response without considering each item in isolation. Many personality tests are ipsative and making participants choose between alternatives can present misleading results and reduce acceptability to candidates.

Paulhaus (1989) delineates socially desirable responding into self-deception (candidates emphasize strong points or talents to themselves) and impression management (candidates act in a way perceived as appropriate for the position). Schmitt and Ryan (1993) argue that job applicants may not deliberately fake responses but may present themselves in a

manner that they interpret as most appropriate to a situation. It is presently unclear however, if tests are taking account of "job desirable behaviour" as well as socially desirable responding (Hough et al, 1990). Schmitt and Ryan (1993) reported that job applicants in their study responded to measures of the Big Five in terms of self-presentations perceived as important for the job in question. This pattern of responding was labelled as the "ideal-employee" factor that is likely to arise from the comparison of perceived self to an idealised image of a typical job incumbent.

Perceived consequences of the test appear to alter the factor structure of responses and image of candidates that emerges, thus, with a five-factor model of personality, the ideal employee responses include many of the items from the "conscientiousness" scale. If the ideal employee constellation of traits corresponds most closely to job performance (or likelihood of being selected) this suggests that the results of Barrick and Mount (1991) could be artifactual.

Cook (1993) proposes several techniques to discourage faking including establishing rapport with the subject, the selection personnel can persuade applicants that doing a job they are not suited to will make them unhappy and unsuccessful. In military studies, subjects have been warned that faking will be severely punished. Items can be made to be subtler and less transparent. Control keys or lie scales/social desirability scales can be included. Forced-choice format using pairs of items that are equated for social desirability may be used though this method is often disliked and creates inter-dependence through scales. Using factual or objective information in items may help to reduce the problem of socially desirable responding. However, the item-designer's assumptions may differ from those of the respondent.

Use of forced-choice formats in item responses is problematic. Expressing interest or similarity in one item necessarily leads to lower scores on other dimensions. This does not allow an individual to express interest in two or more dimensions simultaneously. Johnson, Wood and Blinkhorn (1988) refer to this problem as inter-dependence and assert that it is impossible to use such scales normatively.

Stokes, Saunders, and Owens (1986) tested the coefficient of stability of a biodata instrument over five years and asked external observers to rate the accuracy of responses of participants in verifiable items. Stokes et al report that objective biodata items show the greatest stability, possibly as a result of greater accuracy of responses to objective and moderately subjective factors. Objective items were answered with greater consistency than less objective items. Stokes et al. conclude that socially desirable responding adversely affects item reliability of subjective biodata items.

Several authors have made suggestions for improving reliability of responses. Table 3.3 outlines those of Owens, Glennon and Albright (1966), Schrader and Osburn (1977), Drakeley (1988), and Mael (1991).

Table 3.3 Suggestions for improving the reliability of biodata items

Suggestions for biodata development
Items should be brief and comprehensible
Response choices should be graded on a numerical continuum
An 'escape' option should be offered for alternative responses
All items should have a neutral or equally pleasant tone
Historical and verifiable where possible
Instructions to include a warning that information will be verified and response distortion carries a penalty
Scoring keys should be developed on applicants rather than current employees
Continuous response items should be reduced to a small number of broad response categories
Sensitive items may suffer less distortion but could be useful to control for socially desirable responses.

Lautenschlager & Atwater (1986) report that socially desirable responding or 'faking good' differs across individuals because of differences in the transparency of biodata items. Selection of the 'best' response in biodata is not always easy, especially where items have been designed with a neutral tone.

Several of the reviews have included a plea for psychometric-like lie scales or measures and many have considered response distortion as an individual difference worthy of study in itself (for example, Lautenschlager, 1994; Hogan, 1990). Overall, although response distortion may have consequences in terms of the reliability, construct validity and criterion-related validity of biodata, Drakeley (1988, 1989) concludes that the problems associated with response distortion are overstated although reliability may be improved by following several suggestions when designing biodata items.

3.3 Validity of biodata

Biodata appears to have good criterion-related validity where indices of job performance are used as criteria. Correlation coefficients in the region of 0.35 - 0.45 are commonly reported (for example, Owens, 1976; Mumford & Owens, 1987). These validity coefficients compare well with mental ability tests for the prediction of training success, and turnover (e.g., Hunter & Hunter, 1984; Drakeley, 1988). Asher (1972) asserts that this relative efficacy is due to the degree of overlap between the predictor and criterion in biodata models: In Wernimont and Campbell's (1968) terms, samples of behaviour included in the biodata questionnaire overlap with later samples of behaviour in terms of job performance examples (in comparison to the gap between 'signs' of behaviour such a personality inventory items and examples of job performance behaviour). Mitchell (1996)

asserts that this relative validity arises from the ability of biographical questionnaires to capture typical and maximum performance (psychometrics usually only measure maximum performance in mental ability tests), 'biodata captures both ability and motivation to perform" (Mitchell, 1996:2). Van Zwanenberg and Wilkinson (1996) extend these arguments and assert that biodata is comparatively valid because of the multidimensional nature of the job performance criterion. Therefore, good performance may be achieved in a variety of ways. As a multidimensional predictor, biodata has good predictive validity, commonly containing samples of a wide range of behaviours and proxy measures of general intelligence. Table 3.4 presents the typical validity coefficients reported for biodata measures and job performance ratings for managerial and other occupations.

Table 3.4 The criterion -related validity of biodata

Research	Mean Validity Coefficient
Asher (1972)	.30 - .40
Dunnette (1972)	.34
Reilly & Chao (1982)	.40
Schmidt et al (1982)	.32
Hunter & Hunter (1984)	.37
Barge & Hough (1986)	.35
Mumford & Owens (1987)	.35
Bleisener (1996)	.42 (observed validity) .34 (net validity)
Mitchell (1996)	.30 - .40

Collecting the biodata information and criterion data simultaneously and correlating the scores may measure criterion-related validity concurrently. Predictive validity studies involve the collection of biodata typically from a selected sample of people and correlating it with a criterion (e.g., an index of performance) that is collected after a period of employment. Both forms of criterion-related validity involve methodological problems.

Firstly, the use of correlational analyses to explore the relationship between biodata items and performance outcomes assumes that such relationships are linear. Secondly, the problem of the 'ideal employee' frame of reference that may characterize job applicants may not be present in the developmental pure-research sample. Although predictive studies of criterion-related validity may address some of these issues, restriction of range may be a concern. The developmental sample of a predictive validity study will necessarily consist solely of successful job applicants (present job incumbents).

Overall, the evidence for the criterion-related validity of biodata is good in comparison with other selection methods. However, Hunter and Hunter (1984) have argued that biodata may actually predict "suitability" for an organisational culture rather than actual work efficacy. Hunter and Hunter (1984) employed meta-analytic techniques to re-analyse Reilly and Chao's (1982) data. They report that biodata is a fairly good predictor of performance, but question the utility of biodata in comparison to use of mental ability psychometric tests. Hunter and Hunter conclude that best predictors of entry-level job performance (as assessed by supervisor ratings) are mental ability tests (0.53), job tryouts (0.44) and biodata (0.37). Similarly, Schmitt, Gooding, Noe and Kirsch (1984) report that biodata is not as good a predictor of job performance as assessment centres, but is as good as work samples and mental ability tests.

The variance between validity coefficients from different studies has been attributed to inadequate designs for validation; risk of capitalising on chance; problems of specific item collection and weighting; the impact of varying cross-validation techniques; and the lack of a consensual definition of 'biodata' (Bliesener, 1996).

3.3.1 Face validity: Applicant reactions to selection methods

Participant perception of the fairness of selection methods is closely related to the outcome of the methods (whether they are selected for further testing or not) and the context in which the tests are taken (Elkins & Phillips, 2000; Horvath, Ryan, & Stierwalt, 2000).

Biodata items that do not appear to be relevant to the job may be considered to be invasive. Drakeley (1989) recommends avoiding items of questionable face validity in biodata questionnaires until use becomes more widespread. "The Bureau of National Affairs survey of personnel specialists found that only 4% of respondents used biodata, with perceived invasion of applicants' privacy seen as one of the major reasons for avoiding biodata" (Mael, Connerly, & Morath, 1996:613). Mael et al evaluated biodata items for their perceived invasiveness for a sample of professional and non-professional workers and college students. They report that invasiveness of items is associated with unverifiable, less transparent items with low face validity. Items relating to recall of traumatic events, intimacy, religion, and items resulting in possible stigmatisation were considered most invasive. Mael et al (1996) argue that biodata questionnaires containing items that are non-verifiable, non-controllable, negative events or items casting the respondent in a negative light and items that lack transparency may be deemed irrelevant and unnecessary and items which are personal or intimate and relating to behaviour which would not be demonstrated in the workplace may be perceived as invasive. Mael et al conclude that in terms of avoiding negative reactions to biodata and lowered perceptions of invasiveness, biodata questionnaires should avoid items that are personal, questions about religion and political affiliations, intimate behaviour, and non-controllable events

from personal history. However, concerns about faking limit the possibilities for including transparent items only.

Kandel (1990) reports that biodata items may be unfairly or indefensibly intrusive where they can be said to relate to defamation, invasion of privacy, infliction of emotional distress and / or oppose legislation concerning equal opportunities and the illegality of discrimination on the grounds of race, sex, age, health, or disability (Sharf, 1994: 351). Scharf (1994) asserts that classification or personal data (e.g. marital status) has the greatest potential for being perceived as invasive. Scharf goes on to argue that the apparent criterion-related validity of such items does not sufficiently justify their inclusion in developing biodata inventories and cites the American Psychological Association position of requiring "logically and empirically justified life history questions" (Sharf, 1994:369). Sharf asserts that such guidelines lead to the development of the required theory of underlying relations between job performance behaviours and biodata items contributing to productivity.

Research into models of privacy (e.g. Stone and Stone, 1990) suggests that the purpose of the request for information is an important determinant of perceptions of invasions of privacy. Stone and Stone (1990) report that information gathered for research purposes only is less likely to be viewed as invasive than data collected for personnel records or selection information. Stone and Jones (1997) offer the following suggestions for reducing invasiveness of biodata items:

1. Items directly related to job performance.
2. Items chosen after job analysis (Pace & Schoenfeldt, 1977)

3. Avoiding items that may discriminate between groups of population (e.g. family background, arrests) (Privacy protection Study Commission, 1977).

The criticism that biodata presents a static or 'fatalistic' picture of candidates can be challenged by the use of present life items. Kleiman and Faley (1990) report that biodata items with present-oriented time frame show equal or superior validity to past-oriented items. Present-oriented items may avoid discrimination on the basis of inequality of access to opportunities to demonstrate the behaviours tapped by certain types of biodata item and takes account of recent behaviours and controllable experiences (Kleiman & Faley, 1990; Wilkinson, 1995).

Mael, Connerly & Morath (1999) consider the issue of invasiveness of selection questions from the perspectives of both the organisation and the candidate with particular emphasis on the view of potential co-workers. They argue that fellow employees seek information relating to job-related, controllable, negative and interpersonal information relating to job candidates.

It is important to point out that construct-oriented biodata may be considerably less invasive than personality inventory items and the transparency of items as predictors of work efficacy may lead to increased face validity than traditional empirically derived biographical data questionnaires and traditional psychometric test measures.

3.3.2 Discrimination and legality issues

Where a selection method presents differential validity for different sections of the population it can be said to have adverse impact. The majority of the research indicates that biodata has not created adverse impact in terms of racial minority groups (Cascio, 1976; Owens, 1976; Reilly & Chao, 1982).

Reilly and Chao (1982) present findings that biodata did not discriminate by race but may discriminate against female applicants although Ritchie & Boehm (1977) developed a biodata inventory for AT&T managers which achieved equally good cross-validity for women and men. Criterion indices such as length of employment (perhaps also salary and number of subordinates depending on the limits of the 'glass ceiling' in an organisation) are male biased in respect to female participation in the labour force and possibly in the domestic sphere.

Items concerning socio-economic status pervade biodata (e.g. asking for information about home ownership, parental occupations, etc.). Such items may discriminate against racial minorities and working class applicants (applying for traditionally middle class jobs). However, Drakeley, Herriot, and Jones (1988) report no evidence of differential validity for the prediction of turnover among participants of differing socio-economic and educational backgrounds. Sharf (1994) argues that biodata could be used in place of cognitive ability tests as biodata typically demonstrates comparable predictive validity with less adverse impact.

3.4 Limitations and concerns for users of biodata

Biodata has been criticised for being organisationally specific (e.g. Reilly & Chao, 1982). The potential for lack of generalisability is especially true for empirically derived biodata as such biodata predictors may be sensitive to situational-specific or idiosyncratic variables of the organisation the biodata is developed in (Hunter & Hunter, 1984; Drakeley, 1988, 1989; Owens and Schoenfeldt, 1989; Rothstein, Schmidt, Erwin, Owens, & Sparks, 1990). Laurent (1970) reports positive findings for the generalisability of biodata validity. Laurent found that a biodata instrument developed for the prediction of managerial effectiveness developed in New Jersey could be applied in Norway, Denmark and The Netherlands. Similarly, Campbell, Dunnette, Lawler, and Weick (1970) conducted a study of the Early Identification of Management Potential (EIMP) biodata inventory and uncovered a 'common core' of management skills. This instrument demonstrates generalisable validity across the Standard Oil of New Jersey and affiliate companies in the US. It seems that a person-oriented perspective rather than a job-performance perspective, affords greater generalisability and may hold the key to understanding the processes of individual differences and job performance. Additionally, the Supervisory Profile Record (Richardson, Bellows, Henry & Co. Inc., 1981) derives from 39 organisations and proves to have highly generalisable validity, unaffected by organisation, sex, race, supervisory experience, social class, or education. Rothstein et al (1990) conclude that the results obtained challenge the concept of biodata as organisation-specific when biodata items are developed upon participant samples and job descriptions from multiple organisations. Schmidt and Rothstein (1994) conclude that large sample sizes, data collection from multiple organisations, and cross-organisational keying of the biodata can result in improved generalisability. Carlson, Scullen, Schmidt, Rothstein and

Erwin (1999) report generalisable cross-organisational validity for the biodata component of the Manager Profile Record (MPR) supporting earlier findings of generalisable validity for the MPR across 7334 managers in 24 organisations and using promotional progress as the criterion.

Another criticism of biodata is related to the observation that the criterion-related validity of biodata appears to decay over time (e.g. Thayer, 1977). Items must be rewritten every few years and require large samples for standardization. Item validity may decay due to changes in the nature of work, performance appraisal and the organisational culture. Gunter, Furnham, and Drakeley (1993) argue that certain types of biodata items decay at a faster rate than others. It is noted that unverifiable, subjective items (tapping self-esteem, personal opinions etc.) are potentially more likely to change over time and harder to assess. The criterion used may also affect the length of time an inventory remains valid. For instance, Gunter et al (1993) point out that turnover is especially sensitive to changes in labour market conditions. The consistent conclusion is that biodata should be re-validated every three to five years (e.g. Cascio, 1991; Gunter et al., 1993). In addition, biodata is costly and time consuming to develop, requiring large standardisation samples and the self-report nature of biodata questionnaires falls prey to the response distortion debate outlined in section 2.2.1. (Gunter, Furnham, & Drakeley (1993) offer a review).

Drakeley (1988), Mumford (1996), and others have asserted that the traditionally atheoretical nature of biodata means that understanding of the processes underlying biodata prediction is lost. Similarly, according to Smith (1994) biodata offers moderate coverage of the universal, occupational aspects of work performance but offers no substantial coverage of the relational factors involved.

3.5 The incremental validity of construct oriented biodata

Incremental validity refers to any additional predictive validity that one approach to measurement or prediction may have over another. In this research (biodata for managerial selection) it refers specifically to any additional validity that may be offered by biodata analogue models over traditional psychometric measures of constructs under investigation.

Schmidt and Hunter (1998) present findings from a meta-analytic study of selection methods over 85 years of research. General mental ability is reported to increase the validity (and therefore, the cost-effectiveness or utility) of all selection methods. This research involves observing the predictive validity of selection methods alone (for the most part supervisor ratings of job performance were used as the criterion) and then with general mental ability added. For the prediction of professional or managerial performance, general mental ability (GMA) has a reported validity coefficient of 0.58. Schmidt and Hunter report general (across a range of jobs) coefficients in the region of 0.63 - 0.65 for integrity tests, work sample tests and structured interviews in combination with general mental ability measures.

The incremental validity of selection methods in combination with GMA measures was ascertained in this research by a comparison of validity coefficients before and after the GMA measure was included in a regression analysis. Schmidt and Hunter assert that incremental validity is likely to be greatest where the selection method does not assess the same constructs tapped by GMA tests. The addition of the GMA measure would then

provide a fuller picture of likely performance. This argument is a feature of research by Goldstein, Yusko, Braverman, Smith, and Chung (1998) on the incremental validity of GMA measures and assessment centre exercise predictions of job performance and training ability. Goldstein et al. highlight the racial subgroup differences that emerge out of the introduction of standard GMA tests in the selection procedure. An incremental validity study was conducted to investigate how the validity of separate assessment centre exercises would decrease if GMA were partialled out of the analyses. It appeared that although GMA provides additional validity not captured in assessment centre exercises alone, four out of five exercises remain significant when GMA was removed from the analyses. In particular, the removal of GMA from the 'in basket' task analysis depleted the predictive validity of this component of the assessment centre. Tasks demonstrating lower correlations with the measure of GMA ('subordinate meeting', 'in-basket coaching', 'project presentation' and 'team preparation') remained significant.

Goffin, Rothstein and Johnston (1996) assess the incremental validity for an assessment centre incorporating personality dimensions into predictors of success for candidates of managerial selection. Hierarchical regression analyses were performed to observe the extent to which personality test scores and assessment centre dimensions predicted performance and promotability criteria. Goffin et al conclude that personality testing and assessment centre ratings demonstrate incremental validity over one another when used for prediction of total performance ratings (but not for promotability). They conclude that because personality testing and assessment centres capture different but valid characteristics, combined use should improve prediction.

Vinchur, Schippmann, Switzer and Roth (1998) present a meta-analysis of selection technique validity for sales personnel. Sales ability tests, biodata and interest inventories and the personality dimensions of 'potency' (a component of extroversion) and 'achievement' offered the greatest prediction of performance. Vinchur et al note the incremental validity obtained when GMA is added to sales ability to form a combined predictor of performance (r rises from .26 for sales ability alone, to .36 for sales ability + GMA measure).

Construct-oriented biodata may capture additional variance not assessed by traditional psychometric tests because biodata tends to be designed in such a way as to capture 'samples' rather than 'signs' of behaviour (Wernimont & Campbell, 1968). This means that biodata will offer better prediction of performance because it relates to previous performance rather than intended/perceived likely future performance. Biodata has traditionally been designed to include verifiable historical information, a pre-selection snapshot of likely performance at a work sample test / aptitude for work test.

Following on from Mael and Hirsch (1993) a number of studies have documented the incremental validity of biodata measures over cognitive ability for the prediction of performance, training success, and attrition rates (Allworth & Hesketh, 1998; Dean & Russell, Karas & West, 1998; Mael & Ashforth, 1995).

McMannus and Kelly (1999) extend the work of Mael and Hirsch (1993) in a study of biodata predictors and measures of the five-factor model of personality for insurance agents. McManus and Kelly report that biodata and personality measures may provide incremental validity over each other in predictions of contextual job performance. These

findings lend support to the view that biodata items capturing personality constructs may also capture additional variance.

Mount, Witt and Barrick (2000) extend the findings of Mael and Hirsch (1993) and McMannus and Kelly (1999) in a study of inductively developed biodata scales and measures of both personality items (utilising the five-factor model of personality) and general mental ability. Mount et al use a job-analysis procedure to uncover four components of clerical workers performance. Biodata scales were then developed to include items relating to each component as follows: quality and quantity of work; problem solving ability; interpersonal facilitation and relationships; and retention probability. The biodata scales were composed of items developed on the basis of a rational or 'intuitive' link with the criterion. These items were then keyed against the criterion, so that items were retained in the final biodata scale on an empirical basis and were then cross-validated on a similar sample of job incumbents. The scales were compared with a psychometric measure of personality and a measure of vocabulary, arithmetic and spatial reasoning. Following hierarchical regression in which the effects of length in the job were controlled for, the biodata scales for quality and quantity of work, interpersonal facilitation and retention probability demonstrated significant incremental validity over the psychometric measures of the individual differences constructs. In summary, the incremental validity of biodata scales developed on both an empirical and a rational basis has been quite widely demonstrated through the last decade.

Chapter 4 The nature and antecedents of managerial work

This chapter outlines a number of different approaches attempting to define the nature of managerial work. Task oriented list approaches, competency models, and individual difference approaches will be discussed. Performance criteria must be developed following assessment of what a job involves with regard to the relative importance of different behaviours / competencies that may contribute to effective performance.

4.1 Task-oriented list approaches

Previous research has attempted to outline managerial functions in terms of the roles and or tasks that managers fulfil. In terms of work activities and behaviours, the majority of previous research reports results from small sample investigations, for example, Stewart's diary method research (1967, 1972, & 1975). This 'classical' approach to defining managerial work in terms of how managers spend their time is perhaps most commonly related to the work of Fayol (1949). Fayol's approach involves a move towards description of management functions as distinct from other activities and forms of work. Fayol concludes that managers 'forecast, plan, organise, command, co-ordinate and control the workplace for the enhancement of organisational productivity'. This classical view of managers as rational planners and controllers has come under criticism by writers such as Mintzberg (1975) and Kotter, (1982). Observational research has shown that managers do not spend the majority of their time engaged in rational planning and reflection and that the nature of managerial work is characterised by variety, spontaneity, brevity and interpersonal contact (Mintzberg, 1975).

Tornow and Pinto (1976) also provide a task-oriented classification system of managerial work. The Management Position Questionnaire was administered to managers of different levels and industries, across six large companies. Several categories relating to product marketing and financial strategic planning; co-ordination of other organisational units and personnel, internal business control; products and services responsibility; public and customer relations; advanced consulting; autonomy of action; approval of financial commitments; staff service; supervision; complexity and stress; advanced financial responsibility; and broad personal responsibility emerged following factor analysis. Ten of these factors are in common with the work of Hemphill (1960). Mitchell (1978) presents the following summary categories: planning/scheduling, processing information and ideas, making judgement decisions and problem solving, communicating, both written and orally, interpersonal activity / relations with associates, technical activities responsibility, personal development, personal qualities.

Kotter (1982) provides an in-depth study of 15 general managers between 1976 and 1981. The aim of this research was to uncover the nature of managerial work and the characteristics of the successful general manager. The job demands that emerged are as follows:

- Setting goals
- Balancing resource allocation
- Controlling activities
- Getting information and co-operation from senior management
- Getting staff to co-operate
- Motivating, coordinating and controlling subordinates.

Although the traditional focus has been upon managerial activity, few studies examine managerial behaviour in the context of performance, i.e. with reference to an evaluative component of behaviour. A notable exception to this is the work of Williams (1956).

Williams collected more than 3,500 incidents of executive behaviour from 742 executives across industries, organisations, and geographic locations. Williams (1956) dimensions are planning, organising, and execution of policy, relations with associates, technical competence, co-ordination and integration of activities, work habits and adjustment to the job. Mahoney, Jerdee, & Carroll (1963, 1965) consider that managerial time is allocated into eight basic functions, which can be called 'PRINCESS' factors; Planning, Representing, Investigating, Negotiating, Co-ordinating, Evaluating, Supervising, and Staffing. Mahoney et al expand Fayol's (1949) five functions into eight - and assert that 'representing the company to outside groups constitutes an important managerial function not recognised by Classical theory. Penfield (1975) replicated the findings of Mahoney et al. Table 4.1 outlines the classical functions of management from the research to date.

In summary, the managers did not perform the classical functions of management in an observable or formal sense. Kotter (1982) notes that the behaviour of each manager was shaped by characteristics of the job and of the individual. Kotter asserts that management theory must include at least four classes of variable:

1. Individual (Personal and background information)
2. Contextual (job, business, organisation)
3. Behavioural
4. Effectiveness (performance indicators)

(Kotter, 1982:147-8).

In addition to the four factors listed above, the issue of social or interpersonal interaction is present in a number of the functions lists presented in table 4.3.

Table 4.1 Literature review of Managerial roles / performance dimensions

	Fayol (1949)	Williams (1956)	Hemphill (1959)	Mahoney et al (1963, 1965)	Tornow & Pinto (1976)	Mitchell (1978)	Sternberg (1986)	Kotter (1987)	Yukl (1987)	Ridgeway (1992)	Borman & Brush (1993)	Watson (1994)	New (1996)
Planning	*	*	*	*		*	*	*	*	*	*	*	*
Supervision /influence	*		*	*	*		*	*	*		*	*	
Organising	*	*			*		*		*			*	*
Motivation								*	*				*
Creativity										*			*
Co-ordination	*	*		*	*			*	*			*	*
Control	*		*		*		*	*					
Technical competence		*	*			*					*		
Relationships & Negotiation	*	*	*	*	*	*	*	*	*		*		
Communication			*			*		*	*		*		
Problem solving				*	*	*	*		*				
Resources			*	*	*			*				*	
Responsibility				*		*						*	
Personal skills			*		*	*	*	*	*		*		
Representation of the organisation				*									

4.2 Competency approaches

New (1996) defines competency as 'an observable activity directed towards accomplishing a task. Expressed simply it is the way we do our work.' (1996:45).

The work of Boyatzis (1982) delineates 12 characteristics presumed necessary or important for success of managers. The aim of the research was to introduce the characteristics and activities into management education with the focus strongly placed upon the teaching of practical skills and abilities (Watson, 1994). Boyatzis (1982)

recommends that job demands, organisational environment demands, and the individual attributes of the person be considered in any model of management.

Table 4.2 Competency approach and work clusters (From Boyatzis, 1982)

Cluster	Competencies
Goal action & Management	Planning, organising,
Leadership	Planning
Human Resources Management	Organising
Directing Subordinates	Controlling
Focus on others	Relationship building
Specialized knowledge	

Hales (1986) asked UK managers to describe the nature of their work. The results suggest (in accordance with earlier task list approaches) that managerial work is fragmented and diverse. Current managerial skill languages include reference to clusters of skills or antecedents of managerial performance including: previous job experience, qualifications, outside interests, personal circumstances, job performance and target achievements, job specific knowledge, specified skills or knowledge, conceptual skills, intelligence, analysis, decision making, judgement, problem solving.

Problems identifying managerial work lead to even greater problems in the identification of competent or effective managerial work (Hales, 1986). Managerial effectiveness can be conceptualised as congruence between what is expected of a manager and the performance of those functions, and the meeting or exceeding of those expectations. Taxonomies of managerial competence, or competencies, have been numerous, table 4.3 presents a range of findings.

Borman & Brush (1993) offer an example of a standard competency approach. They attempted to identify managerial performance dimensions based upon the job activities of

26 managerial job functions across a number of organisations. 246 task dimensions were isolated. The 'mega-dimensions' of Borman and Brush 'represents the content of many behavioural examples and activities from managerial jobs in a wide range of organisations' (1993:18). It is difficult to argue that one or other taxonomy of managerial performance is 'best'. These categories may be 'useful as a benchmark set of categories against which to compare the dimensions emerging from studies of managerial performance for individual jobs or in individual organisations' (1993:19). Four categories containing a total of 18 'mega-dimensions' is then outlined. Managerial work was summarised in relation to tasks relating to:

1. **Interpersonal dealings and communication** (communicating effectively and keeping others informed; representing the organisation to customers and the public; maintaining good working relationships; selling/influencing)

2. **Leadership and supervision** (training, coaching and developing subordinates; guiding, directing and motivating subordinates and providing feedback; co-ordinating subordinates and other resources to get the job done; maintaining good working relationships).

3. **Technical activities and the 'mechanics of management'** (planning and organising; technical proficiency; administration and paperwork; decision making/problem solving staffing, maintaining staff and workforce; monitoring and controlling resources; delegating; collecting and interpreting data).

4. **Useful personal behaviour and skills** persisting to reach goals; handling crises and stress; organisational commitment.

Torrington and Weightman (1994) present similar findings to Kotter in that management agendas for working and networks are seen as playing a crucial role for managerial efficacy. The STAMP mnemonic device was developed to outline the various aspects of managerial work: S-Social, T-Technical, A-Administrative, M-Managerial, P-Personal. Torrington and Weightman view know-how, problem solving, and accountability as the core of management.

Watson (1994) offers a critique of the 'POSDCORB' (*Planning, Organising, Staffing, Directing, Co-ordinating, Reporting, Budgeting)* mnemonic of Gulick (1937) to demonstrate the large effect that classical functions theorists such as Fayol (1949) have had on management education. Such activities are regarded as essential to managerial work and have been incorporated into much of management education following research in the 1980's competency-based approach to management education.

Table 4.3 Managerial Competency / Skill Languages

Bennis & Nanus (1985)	Gardner (1987)	Kouzes & Posner (1987)	Kotter (1988)	McCall (1994)	Sprietzer et al (1997)	McCredie & Shackleton (2000:113)
Compelling vision	Intelligence judgement Understand followers	Envision the future	Keen mind	Analytically agile	Culturally adventurous, takes risks	Action bias Analysis
Alignment Creation of manning	People skills Able to motivate	Enlist others Foster collaboration Build commitment	Strong interpersonal skills	Special talent with people	Courageous	Self-confidence Initiative
	Task competence	Linking rewards with performance	Broad knowledge of industry and company Broad set of relationships	Know the business		
	Capacity to win and hold trust	Strengthens others	Excellent reputation High integrity	Broadly respected	Sensitive to cultural differences	Emotional evenness
	Physical vitality Acceptance of responsibility	Confronting and changing status quo sets the example	High energy level Strong drive to lead	Action oriented	Brings out the best in people	Directiveness Developing relationships
	Adaptability Flexible approach			Resourceful	Insightful Integrity	
Learns form mistakes		Experiment take risks and learn from mistakes		Learn from mistakes	Business knowledge	Broad interests
	Courage, resolution, steadiness, confidence			Courageous	Committed	Control
				Sense of adventure	Flexible	
Dedication		Celebrate accomplishments		Passionate	Seeks and uses feedback, open to criticism	
Know oneself				Open to learning	Seeks learning opportunities	

Rose and Baydoun (1995) conducted a survey of the available literature on the nature of managerial work from 1960 and report that Hemphill's (1960) Executive Position Description Questionnaire provides a thorough description of the functions of managerial work. Hemphill asked managers in five organisations to describe their work in terms of separate tasks. Factor analysis uncovered the job components displayed in table 4.4.

Table 4.4 Managerial job components cited in Rose & Baydoun, (1995)

Results of factor analysis
Providing a staff service in non-operational areas
Supervision of work
Internal business control
Technological aspects with products and markets
Human, community, and social affairs
Long range planning
Exercise of authority
Business reputation
Personal demands
Preservation of assets

Some dimensions of job behaviour are easier to observe and evaluate than others. Managerial work involves decision-making, delegating, communication and administrative competence and leadership. Such dimensions may be harder to evaluate than more concrete tasks (Wohlers & London, 1989). Similarly, some dimensions may be given more weight or evaluated more accurately than others (Borman, 1979).

McCredie and Shackleton (2000) and McCredie (2000) apply Kotter's (1982) two-dimensional role of the general manager in which the importance of lateral relationships and the time span of activities are integral to the model of managerial competence. McCredie and Shackleton investigate 'threshold competencies' differentiating between poor and average performance and competencies distinguishing superior performance for a sample of ninety-two current and former autonomous and semi-autonomous divisional general managers. McCredie and Shackleton support the view of Boyatzis (1982) wherein competencies are viewed as 'generic managerial skills and underlying self-concept scales' (McCredie & Shackleton, 2000:107). The following competency clusters emerged:

Table 4.5 Managerial competencies: The unit general manager

(McCredie & Shackleton, 2000)

Cluster	Competencies
Results orientation	Achievement drive Energy Initiative Control Bias for action
Intellectual abilities	Analysis Broad interests
Interpersonal abilities	Directiveness Developing relationships
Resilience and adaptability	Integrity (measure unavailable) Self-reliance/confidence Emotionally even

The majority of these hypothesised competencies were supported at the skill level and were underpinned by 16PF measures of parallel personality constructs. Table 4.5 illustrates those competencies that have statistical significance as predictors of skills ratings and a rating-based 'index of overall success'.

Hayes, Rose-Quirie & Allinson (2000) support the notion of general competencies and conclude that general managerial characteristics and abilities can provide the framework under which more specific function-relevant attributes and skills may operate.

Mintzberg (1975) argues that the classical view presents a mythological view of managerial work as easily categorised, stable and largely resource-driven. He asserts that in reality managerial work is characterised by fragmentation, brevity and informality. Following an observational study of five Chief Executive Officers, Mintzberg reports evidence that 'managers play a complex and intertwined combination of interpersonal, informational, and decisional roles.' (Mintzberg, 1975:49). Mintzberg summarizes the

skills and roles of the manager, including; development of peer relationships, carrying out negotiations, and motivating subordinates. The functions may condense into three primary roles that are influenced by type of industry, environmental factors and stability and public versus private financial sponsorship:

1. Interpersonal roles - the figure head; the leader; the liaison role
2. Informational roles - monitor, disseminator, spokesman
3. Decisional roles - entrepreneur, disturbance handler, resource allocator, negotiator.

This view of fragmentation of the manager's role is supported by the replication study by Kurke and Aldrichs (1983) where four senior level managers were observed and interviewed over the period of a working week. A smaller replication study (Choran, 1969; cited in Kurke and Aldrich, 1983) showed similar findings.

Stephens (1993) conducted a comparative study between Mintzberg's original data and an observational study of five Chief Information Officers. He notes considerable changes in the work place since Mintzberg's original study was carried out. For example, the five CIO's of this study shared the frantic pace of work with the CEO's of the earlier study but a greater proportion of time was spent in reflection and planning and the IT managers spend less time in verbal communication and more time in a resource allocation role than the CEO's, confirming the influential effects of industry type upon work duties and functions. Noordegraf and Stewart (2000) assert that Mintzberg's 'managerial behaviour approach' should be given fresh attention with the focus upon the integration of disparate private and public sector management and further cross cultural experiences in order to contextualise management theory and research.

The overlap between the 'task list' approaches could be said to suggest that ' a common core of aptitudes may be required for managerial success' (Rose & Baydoun, 1995: 144). However, Stewart (1976, 1989) asserts that the focus upon the similarities in managers' jobs across organisations neglects important differences in managerial tasks and calls for a distinction between the study of managerial work, managerial jobs, and managerial behaviour. Stewart (1991) reports that managerial work varies greatly between managers as a function of hierarchical level, function, individual differences and personal attributes of the managers involved, amount and level of power and discretion held by the manager, time pressures and the cultural context in which the manager works. Further limitations of Mintzberg's theory is the discreteness (or lack of) of the ten identified roles of the manager (McCall & Segrist, 1980), and the argument that Mintzberg's theory lacks specificity and does not adequately point out the relationship between role types and effectiveness as a manager (McCall, 1994). However, Carroll and Gillen (1987) assert that the classical functions (Fayol, 1949) remain useful in describing managerial work. As Watson (1994) has pointed out, the focus of research on observational and self-report studies ignores the unseen (possibly most important) aspects of managerial function (Carroll & Taylor, 1968) such as planning, reflection, and so on, which may underlie and connect apparently fragmented and disorganised work behaviour. Carroll and Gillen argue that conclusions about what activities managers undertake and why are futile until the manager is directly asked about the function of the task or role in relation to his/her work. Carrol and Gillen (1987) present a model of the manager at work showing a distinction between goals and tasks. This model integrates the perspectives of Mintzberg (1975) and Kotter (1982).

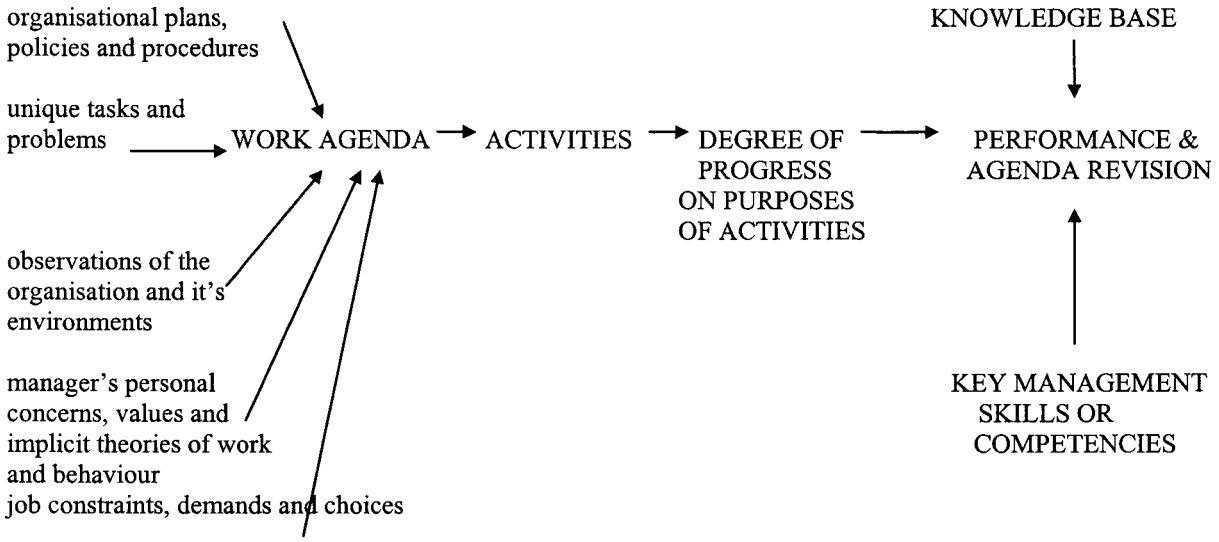


Figure 4.1 Model of the influences upon managerial work (From Carroll and Gillen, 1987: 47)

Pye (1991) argues that the essence of managerial work is tacitly known but cannot be directly outlined with reference to activities. Assessment centres designed to measure managerial qualities may offer global judgements with individual dimension ratings obtained on an ad hoc basis (e.g. Neidig & Neidig, 1984; Sackett & Dreher, 1982). Similarly, Watson (1994) distinguishes 'function' from 'activity' and considers the core management function to be one of steering the organisation toward goals. The basic rationale of management is 'one of establishing and maintaining work organisations as complete entities' (Watson, 1994:38). Managers are envisioned as directors of the organisation first and foremost with specific functional or departmental roles assuming secondary importance next to this goal. Reed (1984) also asserts that the study of management must be sensitive to the context in which the power relations of the manager and subordinates are taking place. He conceptualises management as the unquestioning maintenance of the status quo and therefore as a 'secondary social practice'. Similarly, Rosen (1984) argues that traditional management theory uncritically reproduces the dominant power order that characterises the labour market. This perpetuates the existing

dominance hierarchy and status quo. Rosen asserts that the interdependent relationship between management outcomes and bureaucratic social relations must be recognised within the study of management.

Willmott (1984) offers a related critique of traditional approaches to studying management. Task oriented approaches (focused on outputs of management and activities) are said to present neutral perspectives on the political labour market (in the same way that 'functionalist' accounts offer descriptions of sociological institutions and processes without questioning them). Such task-lists are said to offer decontextualised, depoliticised accounts of managerial influence. Willmott envisages the role of manager as one who is 'institutionally empowered to determine or regulate the activities of others' (Willmott, 1984:350). Willmott (1987) goes on to criticise Mintzberg's (1975) failure to question the origins of the 'formal authority' within which managerial roles are located. Mintzberg's work is considered to embody a 'unitary' standpoint with managerial work simply regarded to be the expression of a division of labour for the achievement of organisational goals.

Kotter's (1982) study has been criticised because it does not adequately examine the political aspect of managerial work and does not extend the observations of managerial networks into the wider institutional and societal context. Kotter's work is considered to present a 'pluralist' perspective with an acknowledgment of the importance of intergroup interaction and networking although this perspective appears to equate power simply with interpersonal relations. Willmott (1984) argues for a 'radical political perspective' that may be 'interpreted simply as providing a demystifying management critique of established accounts of managerial work.' (Willmott, 1984: 364). The 'Radical' account

is said to incorporate the political structures of society within the account of managerial work and status. Managerial work is conceptualised as 'a medium and outcome of the structural properties of a social system founded upon the contradiction between socialized production and private appropriation.' (Willmott, 1987: 268).

Martinko and Gardner (1991) report that the criteria for distinguishing moderate from superior performance among a sample of 41 school principals performance included; minimal competencies, achievement tests, superintendent's rankings of principals, three-year tenure. Although this research supports the literature detailing the brief and fragmented nature of managerial work, there is little support for the proposition that managerial behaviour is related to performance. (Carlson and Mintzberg said that effective managers organise time differently and better than less effective managers.).

Watson (1994) asked the managers in his study what they thought made a manager successful. Willingness to delegate, recognition of others' success, empowerment of staff, having a balance of knowledge of the business and of the people one is working with were the most common phrases reported. Watson concludes that competence may be thought of as the outcome of the cultural context as much as the personal attributes of specific managers.

4.3 Individual difference approaches

Wilkinson (1995) argues that recruitment based upon current competencies is likely to prove 'short sighted' given the increasingly rapid rate of organisational change (Wilkinson, 1995:207). With this in mind a person-oriented or individual differences approach based upon candidate potential is likely to offer greater utility.

Cognitive ability has long been associated with increased ratings of job performance (Ghiselli, 1966, 1971, 1973; Hunter, 1983; Schmidt, Hunter & Outerbridge, 1986; Borman, White, & Dorsey, 1995, Schmidt & Hunter, 1998).

Schmidt and Hunter (1998) present a report of the incremental validity to be gained when cognitive ability tests are combined with existing selection approaches utilising interviews (structured and unstructured), personality tests, assessment centres, references and biodata. Schmidt and Hunter conclude that including a cognitive ability tests in any selection procedure can increase the utility of any selection system producing significant savings in human resource expenditure.

With regard to personality traits and characteristics, Jordan, Herriot, and Chalmers, (1991) support the work of Barrick and Mount (1991) among others in suggesting that performance in some aspects of managerial work may be enhanced by personality traits/aspects. Furthermore, Jordan et al (1993) reported that different personality traits enhance the performance of managers working in different types of managerial function. For example, data processing managers were reported to be more conscientious than others. Finance managers appeared to be more 'astute' than research or development managers (from Smith, 1994:20). Smith (1994) argues that meta-analytic techniques may underestimate the influence of personality traits on job performance because personality cannot be measured as a unitary concept in the same way that cognitive ability can. Current meta-analytic practice often measures the influence of all traits (based on the acceptance of the Big Five model of personality as outlined by Digman, 1990). The

inclusion of 'irrelevant traits' in the calculations will decrease the average correlation achieved for the contribution of personality traits to job performance.

Gulliford (1991) argues that the complexity and breadth of managerial tasks and the elusive nature of managerial work intensifies the usual problems of establishing the correlates of, and criteria for, effective performance. Gulliford presents a "triaxal model" of work performance building upon the work of Blumberg and Pringle (1982) wherein the opportunity, individual capacity and willingness to perform constitute the three fundamental antecedents of work performance. Gulliford's model proposes that aptitude determines ability to perform and personality factors determine the direction and limits of performance. From this perspective it seems wise to take measures or indicators of both mental ability and personality characteristics to provide a wider picture of the applicant. This is a relatively modern perspective and builds upon the research suggesting that personality measures can provide information that is not given by using mental ability tests alone. The author suggests that the arrows to 'opportunity' and 'inclinations' should point towards these terms and also point back up to behaviour in order to demonstrate the reciprocal interaction of the factors listed.

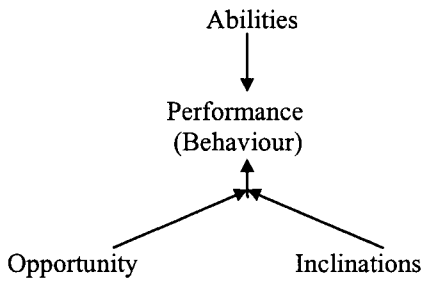


Figure 4.2 Triaxal model of job performance (Gulliford, 1991:29)

Competency lists frequently resort to descriptions of personal characteristics of 'effective managers' possibly as a result of the elusive nature of managerial work and function. Kotter (1982) reports that all of the general managers in his study were highly specialised and knowledgeable about their industry. The majority of them had spent most of their career in one industry although they often moved through jobs and promotions quite quickly. All of the general managers reported liking and having a desire for power and achievement. They were ambitious, emotionally stable (apart from the two entrepreneurs), and optimistic in outlook. Interpersonally, the mangers were personable and good at developing relationships, had wide ranges of interests, above average intelligence (though not brilliant), were moderately strong analytically, and were strong intuitively. All managers had upwardly mobile parents; both original parents at home while growing up; close relationship with one or both parents; at least one parent with college education; father was associated with business or management; all had siblings. All had at least a first degree in business and were student leaders in high school/college or both. All joined an industry that fitted with personal interests and values; (Kotter, 1982, from page 36 and 45). Kotter refers to the profile of the successful general manager as involving a 'success syndrome'. All of the managers moved to new positions rapidly and were upwardly mobile.

Okechuku (1994) studied the personal characteristics associated with managerial efficacy in China, Hong Kong and Canada. In each of the countries self-actualisation, self-assurance, and intellectual ability were considered to be most important.

Similarly, Judge, Cable, Boudreau, and Bretz (1995) investigated the demographic, organisational, and industry/region variables in connection with career success of 1,388

US executives. Criteria of career success included salary, ascendancy, (objective) and job satisfaction, and career satisfaction (subjective). Judge et al report that demographic, human capital, motivational and organisational variables explained variance in objective success and in career satisfaction. Educational level, quality, prestige and degree type predicted financial success. Motivational and organisational variables explained variance in job satisfaction. Preliminary evidence suggests that the variables that lead to objective career success often are quite different from those that lead to subjectively defined success.

In an attempt to identify leadership potential in future international executives, McCall (1994) conducted research that is grounded in a developmental philosophy of management. McCall points out that collecting individual difference information (such as cognitive ability measures, personality data etc.) from existing successful managers may not adequately identify the potential for managerial effectiveness in younger or new-to-the-field applicants. Potential managers may not demonstrate exactly the same qualities or attributes as existing effective managers. It is possible that the qualities of managerial job incumbents are the result of on-the-job experience and organisational cultural mores. McCall defines potential as 'the ability to take advantage of the experiences that will be offered' (1994: 49).

Cortina, Docherty, Schmitt, Kaufman, and Smith (1992) argue that previous research investigating the relationships between personality and job performance has produced mixed results because of the lack of an existing taxonomy for personality dimensions pre-'Big Five'. Writers such as Briggs (1989, 1992) and McAdams (1992) have asserted that the Big Five paradigm is too broad to have any real predictive power. Hough (1992)

asserts that a nine-factor model provides higher predictive validity for job performance. It is likely that specific facets of personality constructs predict performance better than global measures (e.g. affiliation and/or social confidence rather than the trait of extroversion).

Similarly, Schmitt and Ryan (1993) assert that the five-factor model of personality may be limited in the prediction of job performance because of factor divergence across different populations. There is also limited evidence for the five-factor structure among non-volunteer populations in research. These findings clash with those of Mount and Barrick (1998) reporting a consistent relationship between conscientiousness and job performance across occupational category and level. Extroversion is also significantly and positively correlated with managerial and sale personnel performance. The majority of research into the job performance-personality relationship has involved the use of a Big Five model of personality (for example, Barrick and Mount, 1991; Mount & Barrick, 1998) captured by the NEO-PI. The five characteristics have been given separate names by different authors but the usual classification distinguishes surgency or extroversion; emotional stability or neuroticism; agreeableness; conscientiousness; and intellectance, creativity, or openness to experience. However, more recent work by Robertson, Baron, Gibbons, MacIver and Nyfield (2000) contest the idea that conscientiousness is positively associated with enhanced performance in all managerial jobs. They point out that Costa and McCrae (1992) include achievement as a facet of conscientiousness but Hogan (1991) does not. Robertson et al conceptualise conscientiousness as the 'forward planning, conscientiousness and detail conscious' scale of the Occupational Personality Questionnaire. Robertson et al report that conscientiousness was positively and significantly correlated with being 'organised' and 'quality driven' whereas promotability

was associated with 'articulate, decisive, flexible, innovative, motivated, persuasive' (Robertson et al, 2000: 178).

Loevinger (1994) offers a cautionary note on the confusion of the terms conformity and conscience in her critique of the five-factor model. Loevinger asserts that the 'five-factor' model conceptualisation of conscientiousness actually describes the 'conformity' stage of her ego-development theory. In this taxonomy, conformity is characterised by adherence to social rules and norms. This stage precedes a 'conscience' stage of ego development in which moral development and ethical codes for governing behaviour originates within the individual rather than from parents or societal (organisational) rules and values. In this sense, conscientiousness in this account of personality relates to the extent to which the individual is focussed inwards onto their own system of ethical standards rather than a superficial acceptance of imposed roles and social mores. As far as this distinction makes for semantic clarity, the five-factor account of personality cannot distinguish between the two stages. Loevinger's major argument with the five-factor model proposed by Goldberg (1990) and McCrae and Costa (1990) is focussed upon the 'exclusive hegemony' that 'big-five' proponents claim over personality functioning (Loevinger, 1994:3).
Block (1995) presents an equally compelling though strikingly different critique of the five-factor account of personality. Block outlines the criteria for the development of personality theory including:

1. The constructs must be broad enough so that a meaningful consensus may be achieved in understanding each of the component traits or stages of the theory.

2. The constructs should enable predictions about behaviour that are detailed enough to permit a comparison between individuals.

3. The accepted theory must offer the most parsimonious account of the variance observed.

4. Measures of personality must reflect a theoretical basis.

5. Theories of personality and approaches to personality measurement must be formulated on the basis of scientific truth rather than the consideration of efficient and simple measurement. Theory should then inform measurement

On the basis of this approach, Block criticises the five-factor model for a 'bottom-up' approach wherein 'no identifiable hypotheses, theories, or models guided the emergence of or decision on this five-fold space (although some have been offered post hoc).' (Block, 1995:188). Block asserts that the five-factor model offers reasonable descriptive power for the results of the factor analytic studies involved in the development of the model. However, the fact that five-factors may adequately describe the result does not necessarily mean that the factors emerging from the analyses provide a comprehensive account of personality structure and development. The nature of factor analysis is subjective in so far as the emergent factors are labelled. Furthermore, the extent to which a single factor may contribute to variance across statistical results does not automatically qualify that factor as psychologically important. Block concludes that the five-factor approach should not constrain the future of personality psychology and similarly, it should not constrain the future of personnel psychology. On the basis of these cautionary words, the present research project incorporates an attempt to produce explanations of the likely contribution of personality factors to performance variance and calls upon Eysenck's (1957) theory of extroversion and emotional stability in order to do so.

In general, meta-analytic and validity generalisation studies on measures of personality have provided rather disappointing and inconsistent results (e.g. Guion & Gottier, 1965;

Lent, Aurbach, & Levin 1971; Dunnette, 1972; & Ghiselli, 1973; Schmitt, Gooding, Noe and Kirsch, 1984; Blinkthorne & Johnson, 1990). However, several investigations (e.g. Hough, 1988; Barrick and Mount, 1991; Tett, Jackson and Rothstein, 1991) provide grounds for greater optimism for personality as a predictor of job performance. For example, Hogan, Carpenter, Briggs, and Hansson (1985) calculate a mean personality test validity of 0.23 for job proficiency prediction. Although this compares unfavourably with the larger, consistently positive correlations between mental ability and job performance, personality and interest measures appear to show promise for selection purposes. Robertson and Iles (1988) conclude that managerial performance may be predictable from measurement of personality characteristics in certain situations and research paradigms although it is premature to specify a set of traits to be associated with managerial success universally.

In managerial work, extroversion has been associated with enhanced job performance (Bendig, 1963; Barrick & Mount, 1991; Piedmont & Weinstein, 1993). Neuroticism has been chosen for further investigation because as yet the findings are unclear concerning the potential contribution of instability or stability of emotional response to managerial performance. Neuroticism has been positively related to academic achievement in higher education institutions and in other highly selected groups (Entwhistle & Cunningham, 1971; Eysenck, & Eysenck, 1985). Eysenck (1994) argues that neuroticism is likely to correlate positively with achievement among adults but is not expected to correlate highly with measures of IQ. Neuroticism has been negatively associated with managerial interest (Bendig, 1963) and job performance (Hough, Eaton, Dunnette, Kamp, & McCloy, 1990; Piedmont & Weinstein, 1993; Salgado, 1997). Neuroticism did not appear to be significantly associated with managerial performance in the Barrick & Mount (1991)

study, perhaps due to a 'selecting out' of neurotic candidates in the selection process.

Judge and Bono (2000) report the findings of a correlation study between 'transformational leadership' with personality traits measured on the NEO-PI measure of the five-factor model of personality. The results suggest that the traits of agreeableness, extroversion and openness to experience emerged as predictors of transformational leadership. Transformational leadership is said to be associated with leadership effectiveness and can be summarised as the tendency to inspire others to aspire to an idealised vision of the future (Bass, 1985). Judge and Bono point out that the strongest predictor, agreeableness, is likely to correlate positively with team working and group success through the constructs of cooperation and charisma. Neither neuroticism or conscientiousness are related to transformational leadership

4.4 Functional Job Analysis

In contrast with the previous approaches, Fine and Wiley (1955), the objects of worker behaviours can be distinguished as relating work with 'Things', 'Data' or 'People'. Within managerial work, it may be proposed that these categories may relate to aspects of working as follows:

- Things - precision working, setting up, operating-controlling
- Data - synthesising, innovating, co-ordinating
- People- mentoring, negotiating, supervising

Core concepts of FJA include

- A distinction between what workers do (behaviours) versus what gets done (results)
- The objects of work categorised as 'Things', 'Data', 'People'
- Worker instructions (prescription versus discretion)
- Adaptive skills as driver for function and specific content skills
- FJA as a systems approach (linking behaviour, Knowledge Skills and Abilities, with results)

Fine and Getkate (1995) describe functional, holistic, organisational, and premium or special adaptations in order to provide a comprehensive and holistic account of the range of factors influencing each employee in the working context. They present benchmark tasks for a sample of executive managers and executive directors in relation to Things, Data, People, hierarchically organised by level (relative simplicity / complexity) and by orientation (involvement or movement towards a dominance of work with Things, Data, or People).

The orientation of the managers in the research for this thesis was explored by assigning percentages to each of the three functions so the total amounts to 100% (reflecting the performance requirements/orientation) of the task). 'Knowledge', 'Skills', and 'Abilities' (KSA's) associated with each task are delineated through a focus-group methodology and a scale of worker instructions illustrating the degree of autonomy or discretion of action each worker has would complete the standard job analysis procedure.

Fine and Getkate (1995) outline the roles and KSAs required of management workers as follows;

> The executive manager 'facilitates strategic planning sessions / meetings with managers and/or employees, drawing on knowledge of strategic planning

> process, principles and guidelines, current corporate direction and goals, and challenges facing the corporation in future and relying on communication and facilitation skills, active listening, organisational skills, and an ability to establish and maintain rapport in order to elicit the participants' view, determine future goals, and coalesce organisation around the shared organisational vision. He or she will develop or define organisational change strategy, with assistance and support of relevant parties as required, drawing on knowledge of organisational change processes, organisational climate, relevant stake holders, marketplace factors and corporate strategy and mission and relying upon patience, communication, persuasion, and presentation skills in order to acquire buy-in for change from the relevant parties and to create an organisational change strategy. The manager may 'develop / create a corporate mission statement with assistance of external/internal facilitator if necessary, drawing upon knowledge of strategic management procedures and process, corporate challenges and constraints, the current organisational culture and previous experience and relying upon communication and listening skills in order to acquire and define a mission statement that reflects the goals and aspirations of the corporation.

In contrast, the executive director may expect to

> 'Brainstorm / discuss problems affecting the entire agency at retreats and general conferences convened for administrative and executive staff, ask questions, probe for possible approaches to problems such as agency identity and problem definition, drawing on understanding of the needs of the agency and relying on leadership skills and ability to instil confidence and trust in order to establish priorities, explore feasible solutions and develop team approaches. Explore with heads of other community agencies areas of mutual interest affecting the poor and minorities and considering projects warranting mutual co-operation and possibly shared funding, drawing on knowledge of agency and it's potential, community needs and understanding of functions performed by other agencies and relying on interpersonal communication skills in order to develop specific co-operation programmes +/or support for ongoing programmes. Direction comes to the worker in terms of needs (tactical,

> organisational, strategic, financial). Workers must call for staff reports and recommendations concerning methods of dealing with them. They co-ordinate both organisational and technical data in order to make decisions and determinations regarding courses of action (outputs) for major sections (divisions or groups) of the organisation.' (Fine & Getkate 1995:144)

Functional Job Analysis may be adapted to produce a generalisable performance appraisal instrument based upon a priori delineations of job function and task. This approach, in combination with the exploration of individual characteristics and/or general competencies, will be adapted in the present research with the aim of developing a comprehensive assessment and investigation of the likely antecedents of managerial job performance.

4.5 Summary

It appears that the classical management functions of Fayol (1949) of planning, organising, co-ordinating, deciding, and controlling do underpin much of managerial behaviour. It must be recognised that such functions may not be directly visible or immediately obvious through behaviour but it is asserted that these are the primary functions of managers in general. Managers in particular may be said to utilise these function through various means largely dependent upon the contextual and political environment of the organisation and the economy at large. The findings of task-list approaches such as Kotter (1982), McCredie & Shackleton (2000) and Mintzberg (1975) offer interesting insight into the

way managers spend their time and the people they deal with but they offer limited scope for understanding the rationale and aims beneath the overt behaviour.

Chapter 5 Appraising Managerial Work: Obtaining the Criterion

Performance appraisal and evaluation is central to the ethos of human resource management and employee development. Assessment that is valid and reliable is fundamental both in research and the practice of meaningful appraisal and equal opportunities.

Managerial effectiveness has been defined in varying ways, for example Boyatzis (1982) describes efficacy in terms of the functions and outputs of the individual job. This "worker-oriented" perspective has been criticized because it fails to specify what the actual job involves and appears to assume a "universality of management" function without a job or task specification.

5.1 'Objective' criteria

Objective criteria, sometimes referred to as 'personnel data', may include production evaluations such as counts of production, sales and error rates, absence, accident, promotion, salary, and disciplinary action records. Berry and Houston (1993) point out that the focus of these indicators is on the outcome of performance rather than the actual process itself. Often no real attempt is made to uncover the contribution of extraneous factors to an individual's performance (e.g. Cook, 1993). Individual performance may reflect current organisational financial state, the efforts of team members, and external issues such as global and market economy changes. Borman (1991) argues that steps to overcome potential sources of contamination may offer some cause for optimism in the application and search for objective performance indicators. In the case of salary, correction should be made against expected salary as predicted by length of employment (Hulin, 1962). In such instances, objective indicators of performance may have some

validity that subjective ratings do not. Meyer (1987) concludes that performance over time (promotional progress) offers a significant increase in validity and reliability as an index of job performance in comparison with supervisory ratings. This effect may be attributable to the longer period of time required to obtain promotional data for individuals (it may represent a truer picture of performance rather than a "snapshot" image from one or two supervisor ratings), and the possibility that promotional decisions may be based on the judgement of several managers or personnel staff, rather than a solitary supervisor. It is also possible that objective standards of performance (such as production levels etc.) may influence promotional decisions to a greater extent than supervisor ratings. The present research can consider this issue by scoring a 'likelihood of promotion' item and an index of previous promotions and conducting correlation and regression analyses of these measures with the global performance score.

5.2 Performance Ratings

Approximately 75% of research reports indicate that ratings are used as performance criteria with supervisor appraisal being the most common source of performance ratings across a range of occupations (Cascio, 1991; Landy & Farr, 1980). The supervisor is likely to know the person's work and need little preparation in order to answer questions about performance. Evaluation is often seen as part of a superior's job, it is therefore often viewed as acceptable to candidates. Herold, Liden, and Leatherwood (1987) suggest that employees view supervisor evaluations as more useful than those from other sources. However, some employees may not have very much contact with their supervisors, this is especially likely to be the case for managerial, technical, and professional employees.

Table 5.1 Reported uses of ratings in performance data

Study	Results
Guion (1965)	81% of published studies used subjective ratings between 1950 and 1955
Blum & Naylor (1968)	46% of published studies 1960-1965 used judgmental indices of performance
Bernardin & Beatty (1984)	90% of human resource managers used supervisor ratings as a method of obtaining rating data

(N.B. Published data refers to articles published in The Journal of Applied Psychology, and/or Personnel Psychology).

The current trend for 360 degree appraisal wherein each employee is rated by himself/herself, a peer, a subordinate and a supervisor supports the research findings and argument that peer, subordinate and self ratings may offer reliable and valid ratings and also capture slightly different aspects of performance. Mount, Judge, Scullen, Sytsman and Hezlett (1998) collected supervisor, self, subordinate and peer ratings for a sample of 2,350 managers and conclude that low inter-rater reliability was found at all levels of the organisation. In this instance, no distinction was made between supervisor and self-rating variability. Vance, MacCallum, Coovert, and Hedge (1988) used confirmatory factor analysis to analyse the construct validity of multiple job performance measures. Ratings from supervisors, peers, and self-ratings were shown to be equally valid indicators of performance, but it was conceded, "some tasks may be more validly assessed by one source rather than another"(1988:74). Although 360 degree appraisal has been proposed as the 'truest' or most 'reasonable' impression of an individual's performance (e.g. Schrader & Steiner, 1996; Yammarino & Waldman, 1993; Yammarino & Atwater, 1997), the 360 degree method has recently been criticised on the grounds that the relative openness of the ratings (the team members know each other and ratings are going to be

seen by the person being rated and be their manager) may result in greater leniency and investigations into halo error suggests that the dimensions are not clearly distinguishable from each other (Furnham & Stringfield, 1998).

Studies investigating the validity and reliability of self and peer ratings reveal mixed findings. For example, Roadman (1964) reports that peer ratings distinguished middle managers that get promoted from those who did not. The majority of research conducted over the last forty years suggests that peer ratings have good predictive validity across a variety of situations. Reilly and Chao (1982) review peer evaluations for MBA Graduates, managers, life insurance agents, sales staff, and other occupations and conclude that peer nominations predict objective criteria such as graduation, promotion, and job survival, better than supervisor ratings. Mumford (1983) argues that peer assessments are good predictors of managerial job performance when; friendship is present (if making oneself liked is part of the job); where there is consensus among raters; and where colleagues witness the behaviour of the ratee for a large part of the working day. Potential moderators of peer rating validity include the stage of acquaintance or friendship at the time of the rating and the likely consequences of the rating. Although the validity of peer ratings as predictors of job performance has consistently been found to be acceptable in experimental studies (e.g. Roadman, 1964; Hollander, 1965; Kane & Lawler, 1978; Landy & Farr, 1980; Reilly & Chao, 1982), and to be related to higher levels of inter-rater agreement (e.g. Springer, 1953), the implementation of this requires caution. Previous research has suggested that peer ratings may show greater leniency than supervisor ratings (Springer, 1953; Rothaus, Morton, & Hanson, 1965; and Zedeck, Imparato, Krausz, & Oleno, 1974). Peer appraisals have also been reported to have greater accuracy when the ratee is perceived to be similar to the rater and may be influenced by personality

characteristics such as trustworthiness and dependability (Borman, White, & Dorsey, 1995). Dobson (1989) concludes that peer assessments can be meaningful in assessment centre contexts providing the appropriate behaviours have been elicited by group exercises. The validity of the ratings is enhanced if there is good correspondence between the nature of assessments and that of the criterion. Peer ratings will not be collected in the present study, as there is growing evidence that candidates do not favour this method of obtaining ratings and given the sample size it is not practical to begin to identify friendship links and co-working patterns. Another possibility is that peer ratings operate along stereotypic perceptions of what makes a good manager rather than actual work performance and ability, that is, peer nominations and ratings may act in a similar way to assessor ratings in assessment centres; i.e. they may reflect the rater's idea of what a good manager is rather than actual managerial performance, although the impressive validity coefficients obtained in research make this conclusion questionable.

The available evidence for the validity of self-assessments in selection presents a fairly discouraging picture (e.g. Holzbach, 1978; Mabe & West, 1982), although few of the studies that have been reviewed actually reviewed self-assessments in a selection context. Hence it is impossible to account for the likely effects of increased motivation in such contexts. Thornton (1980) reports that gathering information for selection purposes is likely to lead to inflated statements of qualifications and suitability. The issues of leniency and fakeability of responses and ratings have been considered empirically. Individuals seem to possess overly favourable perceptions of their own competence and may fake a favourable response. For example, Meyer (1980) asked participants to rate themselves in comparison to others in comparable jobs. At least 40% of respondents rated themselves in the top 10%, typically, only 1 or 2 % rated themselves below average.

The degree of congruence between self-appraisal ratings and those obtained from supervisors and peers has been under recent review. Furnham & Stringfield (1998) report that typically such ratings do not correlate well and greater congruence may be obtained in ratings of performance in observable behaviour rather than cognitive behaviour, a potential problem for ratings of managerial performance. Furnham & Stringfield report results that support the previous findings of leniency in self-appraisals but report that supervisor ratings show greater halo effects (wherein ratings on some job performance dimensions are contaminated by ratings on performance on other dimensions) so the candidate appears to be 'all-round good or bad'.

Following the empirical findings summarised above and those of Reilly and Chao (1982) and Hunter and Hunter (1984), self-assessment of performance will not form part of this investigation.

5.3 Appraisal Rating Systems

In this study a six-point Likert scale was developed in order to produce a global rating of overall job performance from 'very poor' (obtaining a score of '1') to 'excellent' (obtaining a score, for example, of '6'). The rationale for the adoption of this approach to scaling is discussed below. This form of 'graphic rating scale' separates job dimensions

out in an attempt to uncover individual strengths and weaknesses across job dimensions. Though this is the most common kind of scale in use (Berry & Houston, 1993), this method has been criticised for the high levels of halo and leniency effect it often produces (Landy & Farr, 1980).

The Behaviourally Anchored Rating Scale (BARS) (Smith & Kendall, 1963) was proposed as a solution to the problem of anchor and dimension ambiguity found in Likert-style rating scales. BARS employs the Critical Incident Technique (Flanagan, 1954) in order to obtain relevant and valid items for tapping work performance. Although BARS is time consuming and costly to develop and implement, the aim of developing a scale that describes performance from effective to ineffective along a range of work behaviours was extremely promising. However, the cost and time involved may mean that BARS has limited utility for smaller organisations, and scale use may also be limited to the context in which the scale was developed. The scale required in the present research must be generalisable across managerial functions and organisations. It has also been suggested that BARS may moderate the rater's frame of reference, directing attention to limited areas of job performance resulting in bias (Berry & Houston, 1993). However, Campbell, Dunnette, Arvey, and Hellervik (1973) report a BARS format obtaining less leniency error, less halo error, and less unaccounted-for variance than other methods of scaled rating. BARS does not always function better than well thought out graphic rating scales, although it is highly job related and the use of critical incident methodology imbues the technique with a degree of content validity. Landy and Farr (1980) conclude that comparison of BARS with other rating scales often make it difficult to justify the expense involved in BARS development.

The Behavioural Observation Scale (BOS) was developed by Latham and Wexley (1977) to assess performance by recording the frequency of specific behaviours. This is not likely to be appropriate for the assessment of managerial performance as the classical functions of management (Fayol, 1949), i.e. controlling, planning, organising etc. may not be visible or well judged in terms of frequency.

Generation of behavioural examples of work may be problematic for managerial work where the instrument is required to measure performance across a number of industries and organisations. Overall, research indicates that a wide variety of procedures are used to obtain performance-rating scores. Greater exposure and familiarity with the rating task is believed to create greater reliability and accuracy of ratings (Borman, 1977; 1978), this argument may be taken as support of the use of supervisory ratings in the present study.

5.4 Rating error in performance appraisal

Judgmental methods necessarily involve subjectivity and therefore may be influenced by irrelevant information and bias from the rater. Research evidence has shown that there are large individual differences in judgmental ability (E.g. Landy & Farr, 1980; Latham, Wexley, & Pursell, 1975). Inadequate information may lead the rater to make errors as a result of not having enough exposure to the ratee in order to provide an accurate rating;

this may be especially true of supervisor ratings and managerial work where autonomy of function may be an important feature of the job.

The possible errors that may arise in subjective performance ratings have been summarised by Berry and Houston (1993:262).

- Recency error (the rater bases the evaluation on the most recent behaviour)
- Halo effect (performance on one dimension influences the ratings of others)
- Logical error (a rater assumes that items are similar and rates them similarly)
- Proximity error (adjacent items are rated in the same way)
- Leniency error (rater gives high ratings to all candidates in a sample)
- Severity error (rater gives all candidates low or harsh ratings)
- Central tendency (average ratings of all ratees)

The halo error refers to the generalisation of an impression of performance in one dimension to ratings of performance on other dimensions. The halo error is most likely to occur when the performance dimension is unfamiliar to the rater, is ill defined, and is difficult to observe (Cooper, 1981). Cooper (1981) suggests methods of reducing the halo effect including increasing rater-ratee familiarity, using multiple raters, rating from current exposure, and obtaining ratings of central, irrelevant categories, (1980:235). However, training raters to avoid halo errors has been reported to reduce the overall accuracy of the ratings because there exists "true halo" representing the extent to which performance on separate dimensions actually does converge.

Evidence concerning the length of experience and training of raters appears to give a mixed impression of the influence of these variables upon the accuracy and reliability of ratings. For example, Borman (1975) investigated the effects of instructions to avoid halo error in managerial performance ratings. He reported that training sessions significantly

reduced the level of halo effect in results though the overall validity of ratings was unaffected. Borman concludes that a short training and practice session may offer much in the way of increased reliability of ratings. Similarly, Latham, Wexley, and Pursell (1975) considered the impact of training managers to minimise rating errors over a six-month period. Latham et al. report that trained participants did not commit errors of perceived similarity to the ratee, first impression errors, contrast effect errors, or halo effects. However, this was only a simulation study and did not employ examples of real job behaviour, therefore the promising results should not be overestimated in terms of their generalisability to the workplace.

'Rater Error Training' (RET) (Bernardin, 1978; Bernardin and Walter, 1977; Latham et al., 1975) trains raters to be aware of the errors they may make in order to avoid making them. This method may direct attention to relevant stimuli, but may reduce accuracy by ignoring true halo. It may also be questioned because the focus here is placed upon errors rather than good rating skills. Likewise, 'Rater Accuracy Training' (RAT) focuses upon achieving the most accurate representation of performance in the ratings. This method was developed by Pulakos (1984) who also compared the efficacy of RET and RAT and concluded that rater accuracy training achieves the most accurate, and least lenient performance ratings; although rater error training had the largest positive effect in reducing halo error. However, Pulakos also reports a decrease in accuracy when RET and RAT methods were combined, possibly from the failure of candidates thus trained to recognise true halo. Overall, he concludes that training should focus less upon specific constraints and consider wider sources of inaccuracy and unreliability in ratings. The need for longitudinal studies of training and rater accuracy over time is also highlighted. Comparisons of the efficacy of accuracy and error training have shown that error

avoidance training is more accurate and reliable than accuracy-only training (Pulakos, 1984). However, accuracy training is more effective when accuracy is judged by comparing the ratings of true scores generated by experts. Sulsky and Balzer (1988) suggest that swapping error training for accuracy training merely exchanges one set of problems for another.

"Frame of Reference" (FOR) training (Hedge and Kavanagh, 1988) attempts to establish a common frame of reference among raters and provide them with a standard for evaluating work behaviours. This may be particularly complex when dealing with managerial behaviour across functions and organisations. McIntyre, Smith and Hassett (1984) consider the effects of frame of reference rater training and conclude that training had a significant effect upon the halo effect recorded in performance ratings and in the overall accuracy of ratings. This conclusion differs from that of Zedeck and Cascio (1982) who conclude that training had no significant effect upon the discriminative ability of raters and the accuracy of ratings. However, in the study by McIntyre et al. (1984) almost 85% of rating accuracy variance was unaccounted for after the effects of training were considered. Supervisors providing performance ratings in this research will not be trained to avoid error and bias. It is important to note that shared biases across raters may enhance the apparent reliability of ratings while lowering the validity. "Establishing the construct validity of the expert ratings is important to ensure that an appropriate standard for comparison is being used." (Sulsky & Balzer, 1988:504). The general lack of a definition and theory of job performance may lead inevitably to reduced validity and reliability of performance ratings across raters. Definitions of job performance couched in organisational terms rather than individual, worker-oriented terms may be less likely to show promising results in terms of the psychometric properties of performance rating

methods.

The purpose of the rating, the racial and sexual composition of the workforce, the economic and sociological environment and conditions of work, and the number of poor and excellent performers in a work team represent some of the possible moderating influences on performance ratings. Zedeck and Cascio (1982) assert that the purpose of the rating (i.e. selection or training and assessment use) moderates ratings to a greater extent than rater training. Harris, Smith, and Champagne (1995) compared research based and administrative purpose for performance appraisal, the latter showed greater leniency. Explanations of the effects of appraisal purpose include the 'accountability' effect.

Sulsky and Balzer (1988) assert that there is no common definition of performance rating accuracy. They argue that the accuracy of performance ratings have been calculated by comparing raters' performance ratings of a number of ratees on several performance dimensions with ratings from "expert raters", (expert ratings were obtained by computing pooled average scores).

5.5 Rater and ratee characteristics

Several attempts have been made to uncover the factors, which are likely to influence or bias performance ratings. Rater and ratee age, gender and ethnicity have all been considered. For example, research investigating the effects of age upon performance ratings suggests that there are no significant effects of the age of ratees upon the rating given by raters. However, Schwab and Heneman (1978) report that older than average participants evaluated older ratees more harshly than younger raters. There was no evidence that the amount of job experience of rater or ratees influenced the favourability

of ratings. It should be borne in mind that the participants of this study appraised written descriptions about work behaviours rather than observations of job performance. Cleveland and Landy (1981) report that older ratees received lower ratings and younger raters gave more favourable ratings to older and young ratees than older raters did. The age effects were small but statistically significant. The influence of age accounted for between one and four percent of rating variance. It is unclear if rating differences were due to actual performance differences or bias among raters. It is likely that the nature of the job would influence the relationship between age and performance appraisal, further research must take account of this factor.

In relation to gender and race, Landy and Farr (1980), Schmidt and Lappin (1980) and Kraiger and Ford (1985) report differential patterns of ratings in favour of one's own ethnic group. Overall, studies of black and white raters and ratees support the conclusion that both races showed greater variability in ratings of members of their own race, this finding is discussed with reference to the possibility that candidates rate members of their own race with more confidence than those of other races, this observation would support the 'outgroup homogeneity effect' discussed in Social Identity Theory (Tajfel, 1978). Kraiger and Ford (1985) highlight moderating factors including changing the composition of the group: higher proportions of ethnic minority group members resulted in lower intra-race favouritism, and the research setting: field studies show greater effect sizes.

Investigations of the effects of gender on performance ratings present mixed results. This effect is most likely moderated by the job that candidates are being evaluated in. Ezell, Odewahn, and Sherman (1982) conducted a survey of public agency managers being

evaluated for promotion. They report that ratings of female candidates were based more on previous work performance and ratings of men were based more on perceived potential.

Bem (1974) argues that the psychologically androgynous manager is likely to enjoy greater success, as s/he is adaptive, flexible and able to utilise traditionally masculine dominance and feminine nurturing and communication skills across a variety of contexts.

Cames, Vinnicombe, and Singh (2001) reported that 36% of a sample of 66 managerial staff at banks in Luxembourg, Sweden and Italy reported the 'successful manager' as androgynous. 32% of the sample said that successful managers were likely to demonstrate masculine characteristics while only 14% argued that traditionally 'feminine' characteristics embodied their perception of a successful manger. Analysis of interview dialogue revealed that the successful manager was perceived as one who is '100% dedicated to the bank, dedicates almost 90% of his life to work...' (Cames et al., 2001:111). Cames et al argue that for female managers in their sample, this total dedication was perceived as strong instrumentality, a traditionally masculine trait. Cames et al conclude that female managers were more likely than male managers to perceive success as a result of masculine traits and behaviour. However, Liu and Wilson (2001) argue that female managers are more likely to be tenacious than their make counterparts, adopting a democratic style of management that is likely to encourage commitment and support from subordinates and peers at work (Liu & Wilson, 2001:172). It is therefore likely that male and female managers may adopt different roles and styles of management to achieve success, the translation of these differences into performance ratings and organisational reward mechanisms is a fascinating area for research but lies beyond the scope of the current research.

To summarise, Borman, White and Dorsey (1995) suggest that in supervisor ratings especially, ratee race and gender appear to have minimal effects on performance ratings (see also Oppler, Campbell, Pulakos, & Borman, 1992; Pulakos, White, Oppler, & Borman, 1989; Cleveland & Landy, 1981; Schwab and Heneman, 1978; Rosen and Jerdee, 1973). Overall rater and ratee race and gender have been found to account for less than 1%-5% of rating variance (Kraiger & Ford, 1985; Pulakos et al., 1989), and ratee age appears to have no significant effect on performance ratings (Bass & Turner; 1973; Cleveland & Landy, 1981).

Wayne and Liden (1995) report results that suggest that supervisor liking of candidates and perceived similarity to ratees significantly influenced results. Bernardin, Cooke and Villanova (2000) report that raters with high levels of the trait of agreeableness and low levels of conscientiousness generate lenient ratings of peers. Perhaps rater personality should be assessed in an occupational context to explore the potential moderating effects of these traits. Overall, a social psychological approach may shed light upon various potential areas for bias and extend the focus of purely sociological or psychological perspectives discussed above.

5.6 Psychological models of the rating process

Cognitive models of the performance rating process: include Cooper (1981), DeNisi, Cafferty and Meglinno (1984), and Landy and Farr (1980). Of interest is the way that raters observe and encode and recall information and the possible influential biases in this process that may result in halo and other errors (Cooper, 1981).

Feldman (1981) envisages the appraisal task as one of cognitive complexity and argues that the categorisation of ratees in this process occurs on an automatic, unconscious level unless decisions become excessively problematic. The automaticity of information processing in this context led to biased recall of ratee performance depending upon contextual and dispositional factors of the rater taking part in the appraisal process. The bias of performance recall is a result of the prototypical category that a rater perceives a ratee as belonging to. Feldman proposes that person perception represents the interaction of sensory perception, cognition, and social behaviour. Categorisation of candidates rests upon their perceived similarity to other, prototypical members of existing categories, for example a concept of an 'effective manager' and a different concept of an 'ineffective manager'. Feldman accepts that the structure and format of the rating method may interact with the psychological process of categorisation, storage and recall to determine the performance rating. In this model, halo error is the inevitable product of the categorisation of ratees and the subsequent recall of behaviour in the appraisal context.

Cardy and Kehoe (1984) consider the influence of information processing characteristics and style of raters and assert that raters high in selective attention ability provided the most reliable and accurate performance ratings. Large individual differences among raters in ability to recall information associated with performance led Cardy and Kehoe to suggest that objective records of performance and personnel data (e.g. absence records, salary increases etc.) should be maintained whenever possible. Similarly, Schneier (1977) postulates that more cognitively complex raters show less halo effect and range restriction in ratings. Causal modelling/path analysis is another method of showing which factors influence ratings. Hunter (1983) conducted path analysis of four civilian and ten military

studies and found that ratee job knowledge-supervisory ratings path coefficient was much larger than the task proficiency-supervisory ratings path. Cognitive ability had an indirect effect through influence on job knowledge. Schmidt, Hunter and Outerbridge (1986) reported supportive results and add that job experience also had an indirect effect on ratings through an influence on job knowledge.

Landy and Farr (1980) propose that the combination of implicit personality theory with a process model of the appraisal process can achieve greater understanding of the moderating influences and effects of and within job performance ratings. Use of objective data and judgmental methods in combination may provide the most reliable and accurate information, (Cardy & Kehoe, 1984). However, Landy and Farr point out that obtaining complete personnel records or objective data can be problematic and restriction of range in such data may be a problem (for example, accident records may hold information for a small proportion of employees, those having the most accidents). The process model of performance appraisal considers the influences of the following major variables,

- Position characteristics
- Organisational characteristics
- Purpose of the rating

Within this framework, the subsystems of cognitive processes of the rater (observation, storage and encoding, recall), and the method used to obtain the rating, contribute to the appraisal outcome. Such a paradigm can offer greater insight into the rating process and illustrate how ratings are developed. This insight can lend itself to the process of establishing the fairness and acceptability, as well as the validity and reliability of performance rating methods.

Conclusions about the influence of psychological variables on the performance rating process are difficult, the use in many studies of small samples, the variability among expert ratings used as criterion measures, and the differential purposes of each individual study mean that generalisations to the work place (and to managerial work in particular) are problematic.

5.7 Summary

Composite measures of performance tend to be based upon economic constructs rather than the behavioural factors underlying job performance (Ghiselli, 1956). In managerial work, a univariate performance measure may betray the possibility that several managers may perform effectively in the same job for different reasons (Otis, 1940). The autonomous nature of management necessitates the consideration of individual attributes in addition to job and organisational characteristics. Ghiselli (1956) goes on to assert that because the performance requirements for work are likely to change over time, a 'dynamic criterion' focus could delineate specific characteristics or abilities that are good predictors of performance over time (and presumably over job requirement changes). The 'criteria of success' may not be stable (Bass & Barrett, 1981; Austin, Humphreys, & Hulin, 1989). In addition to temporal changes in the job, the scope for autonomic action or heightened experience of discretion in behaviour choices, may amplify the criterion problem when applied to managerial work (e.g. Borman, 1991).

Chapter 6 Methodology: The development and validity of construct oriented biodata

This chapter outlines the processes involved in developing biodata analogue models of psychometric constructs and illustrates preliminary evidence of the reliability and validity of models developed using the construct-oriented approach put forward by Wilkinson (1995).

6.1 Rationale for the preliminary study

A cross-validation study of the models developed by Wilkinson (1995) was performed in order to investigate the validity of the construct–oriented approach to item and scale development. A replication of Wilkinson's analogue model development procedures with additional items in the biodata questionnaire served as an opportunity to extend the models and check the new items out with a pilot sample of potential managers in preparation for the incremental validity study with the managerial job incumbents, and to consider the generalisability of Wilkinson's models across two similar samples (after Cronbach, 1990). Formulaic approaches to cross validity estimation tend to over-estimate validity shrinkage of rational models of biodata and underestimate that of empirically derived models (Mitchell & Klimoski, 1986). In this instance, the approach recommended by Mitchell and Klimoski has been adopted whereby biodata models will be validated against two independent samples of participants in order to obtain the most realistic albeit rather conservative estimate of validity shrinkage.

The large developmental samples required for accurate correlational analyses are often proposed as a frequently cited reason for the low rates of use of biodata. England (1971) recommends a minimum of 75 participants for the standardisation of each biodata key and a further sample of 75 participants for the cross-validation sample. However, Devlin, Abrahams and Edwards (1992) assert that groups of around 50 participants may provide accurate measures of cross validity and argue that large samples are no less likely to prevent shrinkage than smaller samples.

6.2 Extension of Wilkinson's biodata questionnaire

A total of 17 items were added to Wilkinson's (1995) original questionnaire with the intention of predicting neuroticism (see Appendix I). Of the item responses, nine significantly predict neuroticism, ten predict extroversion and four predict critical thinking ability. The section of the biodata questionnaire that shows the greatest prediction of the psychometric measures is the section containing self-perceptions.

The items were designed to capture personality traits without being invasive. In all cases where it was possible, the items refer to job-related behaviour and preferences, for example an item about the need for job security was included in an attempt to predict neuroticism. Items about getting recognition for success and risk taking predict neuroticism.

The biodata questionnaire contains categories of information as outlined in table 6.1 (see Wilkinson, 1995). Design of the neuroticism items has mirrored that of the extroversion

and critical thinking ability items (Wilkinson, 1995). Items include dichotomous or single response continuum items; multi-choice, and single choice non-continuum questions. The items have been designed so that each option has a neutral tone to attempt to reduce socially desirable responding.

Table 6.1 Biodata question categories

Biodata item categories
General background information
Education
Work styles
Interests, activities and hobbies
Self- perception

6.3 Psychometric Measures

The current investigation explored the cross validity of the analogue models of critical thinking ability and extroversion, with the addition of neuroticism. This represents a quasi-rational or construct-oriented approach to biodata development, as discussed in chapter three.

The Watson Glaser Critical Thinking Appraisal provides a measure of critical thinking ability or ability to:

> 'define a problem, select information that is relevant to the solution of the problem, recognize stated and unstated assumptions, formulate and select relevant and promising hypotheses, and draw conclusions validly and judge the validity of inferences' (Wilson & Wagner, 1981).

Such abilities are found in the conceptualisation of 'executive intelligence' (Scriven, Savage, & McCredie, 1994).

The Watson Glaser Thinking Appraisal test consists of 80 items distributed through five categories. The five categories are 'designed to measure different though interdependent, aspects of critical thinking' (Watson & Glaser 1991:3-4). The total scores will be used for validation of the biodata analogue models. This test was administered in an untimed session and the majority of participants completed the test within 60 minutes.

The Eysenck Personality Inventory consists of twenty-four questions assessing the extroversion-introversion factor, twenty-four questions assessing neuroticism-stability factor and a lie-scale of nine questions based upon the 'unlikely virtues' model of responding (Eysenck & Eysenck, 1963). Respondents are required to select a 'yes' or 'no' response for each item and the administration of this test was untimed.

6.4 Construction of the biodata analogue models

Wilkinson (1995) employed correlation, regression and factor analytic approaches to develop biodata analogue models of extroversion, critical thinking and vocational interest. His results suggest that regression and factor analytic approaches do not produce enhanced criterion-validity over correlation-based analogues and the validity of analogue models derived from regression analyses is likely to show shrinkage on cross-validation (Wilkinson, 1995:195). In correlation analysis, weighted item models were developed by selecting those items that produced a significant correlation with critical thinking, extroversion or vocational interest score and multiplying the item number by the weight of the correlation plus or minus (depending on the direction of the following item) all other items demonstrating significant correlations with the [psychometric test score. This procedure created a single weighted item variable that could then be correlated with the

original test score. Unit weighted or 'rational' correlation models were derived in the same way but assigned a weight of plus or minus 1 depending on the direction of the relationship rather than using the actual coefficient as a weight.

6.5 Cross validation study

The following section outlines the steps taken in the development of initial biodata analogue models designed to measure neuroticism, critical thinking ability and extroversion and present the procedure for cross-validation of Wilkinson's (1995) analogue models and models developed from a similar sample of 'potential managers'.

6.5.1 Participants

In order to draw sound conclusions about the generalisability of biodata analogue models of psychometric constructs, the sample used for the preliminary research mirrors that of Wilkinson (1995). A total of 243 (102 males, 140 females, 1 undisclosed gender) participants volunteered to take part in the research. This sample consisted of full and part-time business studies students, practicing managers, HRM professionals, and lecturers.

As figure 6.1 shows, the majority of the sample was aged between 20 and 29 years of age (60.5% of the sample).

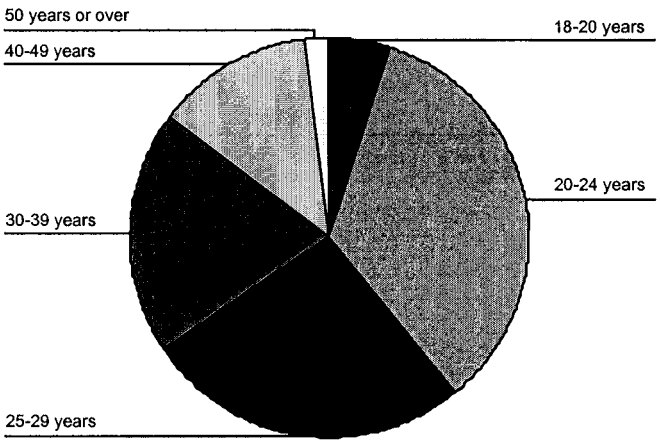


Figure 6.1 Age distribution of the preliminary study sample of 'potential managers'

Eighty per cent of the sample had received their pre-university in England, Wales or Northern Ireland, 8% outside Europe, 6% in the European Community, 4% in Scotland and 1% in Europe (excluding the European Community).

In relation to occupation, 44.4% of the sample was in full time education. The other 135 respondents were either part-time students or engaged in full time managerial or managerial trainee work. Table 6.2 illustrates the occupations of the managers and the part-time students taking part in the research.

Table 6.2 Occupations of part-time students

Management function
General manager
Training / Human Resources
Quality Assurance
Marketing
Administrative
Civil service
Engineering/technical
Lecturer/Higher Education

It can be observed from table 6.2 that the participants cover the majority of typical managerial functions and represent a range of different industry types including the general areas of IT, personnel and human resource, logistics and education.

Table 6.3 Subject areas of student participants

Course topic
Quality Enhancement & Management
Technical Management
Accounts & Financial Management
Public Services Management
Retail Management
Human Resources Management

Table 6.3 illustrates the nature of the course of study engaged in by the student volunteers. This list includes full and part time study as well as undergraduate and postgraduate students. In terms of standards of education, figure 6.2 illustrates the standard of higher education of the sample.

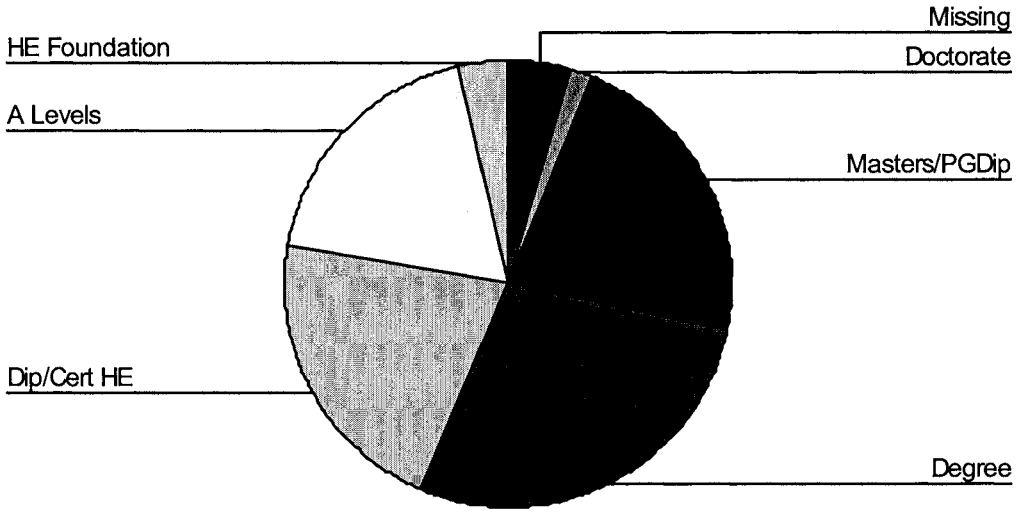


Figure 6.2 Higher education achievements of the sample of 'potential managers'

Almost 25 % of the sample had formal management qualifications and 36 % were members of a learned professional society or institute.

6.5.2 Procedure

Volunteers were recruited from Newcastle Business School to participate in a study on managerial selection. All participants were offered feedback on their scores of critical thinking ability, extroversion and neuroticism. The feedback information is included in Appendix VI. All participants were asked to complete the Biodata questionnaire, the Eysenck Personality Inventory (Form A) and the Watson Glaser Critical Thinking Appraisal (Form C, untimed). The Eysenck Personality Questionnaire has adequate reliabilities (split-half and test-retest reliabilities ranging from 0.74 - 0.91) and substantial validity information and normative data. It is quick to complete and easily administered and scored.

The battery of tests was completed in groups of 25 participants in sessions lasting up to two hours. Response sheets and questionnaires were coded to ensure the confidentiality of recorded data, volunteers were asked to retain their code number in order to collect the correct feedback achieved this. Course leaders and organisational officials maintained a list of participant initials and code numbers in order to prevent the mismatching of feedback with volunteers failing to recall their code numbers. Such information was stored separately from the test scores. Psychometric test feedback was provided on EPI and critical thinking scores for all candidates in written form. One-to-one discussion of test scores was offered to all participants who wished to discuss scores in more detail.

Respondents scoring four or above on the Lie scale of the EPI were excluded from further analysis in an attempt to control for socially desirable responding and ensure the reliability of the biodata analogue models assessing the personality constructs. The psychometric scores obtained are approximately within the ranges described as 'normal' for a student / managerial population (Eysenck & Eysenck, 1963; Watson & Glaser, 1993) though the critical thinking score is slightly lower than that published in the test manual, this is likely to reflect the fact that the Watson Glaser manual norm of 60-61 is the mean score for practising managers rather than trainee managers or students of management. The student sample is slightly more extroverted and neurotic than the managerial sample mean reported by Eysenck but both scores (14 and 11) are within the 'normal' or average range for students. The mean data is presented in table 6.4

Table 6.4 Psychometric test data for the sample of 'Potential Managers'

Variable	Mean	S. Dev.	N	Reported norm for this population
CTA	56	8.19	172	60 - 61
Extroversion	13.93	4.04	242	10 - 11
Neuroticism	10.50	4.74	242	8 - 10

The large standard deviation of the critical thinking scores possibly reflects the contribution of scores from participants who did not have English as their first language. 84% of the sample received their pre-university education in England, Scotland, Wales or Northern Ireland. A further 6% were educated in the European community (excluding the UK), and 8% were educated outside Europe.

Chapter 7 Methodology: The incremental validity of construct-oriented biodata

7.1 The managerial sample

This sample included managers from a medium sized, public sector local transport organisation (n = 114), a private sector air filter company (n= 21), and a private sector manufacturing organisation (n=5). The total sample size was 135. The sample consisted of 118 males, and seventeen females from a range of managerial functions and levels. The total percentage of female managers participating in the study was 12.6%. Table 7.1 outlines frequency data for the sample of managerial workers.

Table 7.1 Descriptive data for the managerial sample

	Mean	Std Dev.	N	Norm for this group
Age	41.76	8.21	134	-
Critical thinking score	59.81	8.56	135	60-61
Extroversion	12.24	4.04	135	10-11
Neuroticism	6.74	4.51	135	8-10

This sample was slightly older than the pilot sample and has a slightly higher average score on the critical thinking appraisal test. The practising managers appeared to be less neurotic than both Mitchell's pilot sample of potential managers and Wilkinson's (1995) sample. This maybe due to the increased neuroticism scores commonly found among samples of student participants rather than being a feature that is specific to the samples of managers (Eysenck & Eysenck, 1963). Alternatively, it is possible that employees demonstrating higher levels of neuroticism are not selected for management or promotion.

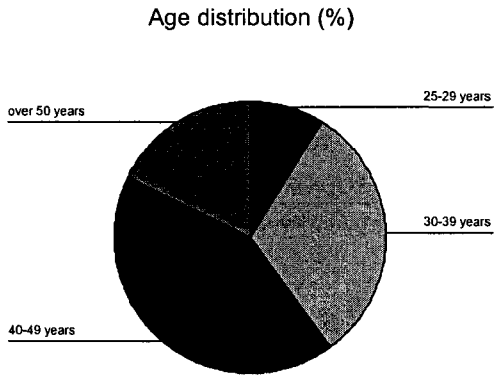


Figure 7.1 Age distribution of managerial sample

7.2 Procedure

All participants were invited to volunteer in the study. All test scores were collected anonymously, using code numbers and were matched against performance data by a 'blind' organisational administrator. Participants were reassured that all scores would be confidential and would not be used for any purpose other than for the biodata project research. All participants were offered oral and written feedback on their psychometric test scores. Appendix III shows the research brief sent out to potential participating organisations.

7.2.1 Psychometric test administration

Data collection proceeded in groups of 25 participants. Following a presentation on the use of biodata and psychometric measurement at work, all participants were invited to complete the 66-item biodata questionnaire (see Appendix II), the Watson Glaser Critical Thinking Appraisal (Form C, untimed), and the Eysenck Personality Inventory. All three tests were completed in a session lasting approximately 90 minutes.

7.3 Development of the performance appraisal instrument

The circular nature of management competency distinction and the managerial person specification, combined with the need to include factors for identifying potential success, necessitates the development of a managerial selection tool that is generalisable and manages to combine indices of previous success with personal attributes and traits. The biodata instrument currently under development can meet each of these requirements in turn.

The performance appraisal instrument (see Appendix VI) obtained ratings of performance from supervisors of the managerial staff and personnel data. The following categories of personnel data were gathered for each participant: absence, salary, number of promotions, and length of time in organisation. Table 7.2 illustrates the average scores for the personnel data.

Table 7.2 Performance data for the managerial sample

	Mean	Std Dev.	Min	Max	N
Total performance rating	4.39	0.88	2	6	120
Salary (k)	21.8	8687	10.4	49.3	134
Promotions	1.65	1.45	0	6	134
Work time (years)	10.86	5.83	0.25	27.00	134
Absence°	1.37	0.69	1	4	134
Career progress*	14.63	12.9	0	60	134

°Absence was rated as indicated on Appendix VI using a scale of '1'(no absence) to '4'(more than 2 months)
** Career progress = promotions / length of time in the organisation*

The concept of 'career progress' (incorporating promotions and length of time at the organisation), may involve more 'typical' levels of performance or life experience than more abstract perceptions of personality consistency sampled by personality

questionnaires and individual boss's subjective ratings of individuals performance.

7.3.1 Functional Job Analysis: A generalisable appraisal instrument

A job analysis model of performance has been adapted to offer a multidimensional approach to appraisal. Campbell, Ford, Rumsey, Pulakos, Borman, Fleker, DeVera, & Reigelhaupt (1990) define job performance as multidimensional 'the total population of behaviours and activities that are judged to be important for accomplishing the goals of the organisation (1990:278) (see also Borman, Hanson & Hedge, 1997; Campbell, 1990, 1991; Conway, 1996; Murphy, 1989, 1996 and chapter four).

This research will involve an overall performance measure and subcategories of performance (see Appendix VI). The combination of objective and subjective rating data reflects the multidimensional nature of managerial work performance and is considered to increase chances of gaining reliable performance data (Cardy & Kehoe, 1984; Bommer, Johnson, Rich, Podsakoff, & Mackenzie, 1995). Job performance represents a combination of individual task performance and organisational behaviour (Campbell, 1991), this proforma represents the performance model of Campbell implicitly (1990, 1991, 1994; Campbell, Ford, Rumsey, Pulakos, Borman, Felker, De Vera, & Riegelhaupt, 1990).

Campbell (1991)

Job performance = [Declarative knowledge (DK) * Procedural knowledge and skills (PKS)* Motivation (M)].

Figure 7.2 Campbell (1991) Model of job performance

Each participant of the present study was appraised individually, in an absolute performance sense rather than in a comparative approach (against the performance of colleagues). Although comparative performance may reduce the need for reliance upon 'notions of abstract standards' of managerial performance, (Wagner & Goffin, 1997), the absolute measure eased data collection and comparison of individuals from different organisations and across departmental boundaries.

The collection of demographic information such as race, gender and age of both the managerial employee and the supervisor providing the performance rating permits the investigation of possible variance from characteristics of the rater and / or ratee.
The inclusion of promotional data in the appraisal tool is based upon the findings of Meyer (1987). Incorporating promotional data in the appraisal instrument may improve the reliability due to the longer period of time required to obtain promotional data for individuals (it may represent a truer picture of performance rather than a "snapshot" image from one or two supervisor ratings), and the possibility that promotional decisions may be based on the judgment of several managers or personnel staff, rather than a solitary supervisor. It is also possible that the influence of objective standards of performance (such as production levels etc.) may influence promotional decisions to a greater extent than supervisor ratings is also discussed.

Borman (1991) considers the 'relevance' of performance criteria. Relevance is defined as the 'correspondence between criteria and actual performance demands of the target job'. The appraisal instrument must avoid both criterion contamination 'when a criterion measure taps variance irrelevant to the performance requirements' and criterion deficiency, 'when a set of criteria for a job fails to measure one or more of its important performance areas' (Borman, 1991:272).

Job knowledge is conceptualised as a result of time in the organisation, intelligence and access to important information and groups or individuals. Effective managerial work may be thought of as job knowledge, intelligence and traits specific to industry or work function. The performance appraisal under development must be broad enough to capture all of the relevant aspects of each manager's work without assessing irrelevant behaviour or characteristics. To this end, Functional Job Analysis (after Fine & Wiley, 1955) has been applied.

The importance of having a generalisable instrument that may be applied across managerial functions, levels and organisations is central to this research. As such, the distinction of task orientation towards work with people, data and things has been adopted for use with broad competency statements relating to each of these areas of work. The in-depth focus group approach is not practical for the validation of the biodata inventory under development though this approach certainly offers scope for future research developments.

7.3.2 Performance ratings

A global measure of overall job performance was collected. This univariate 'total performance' criterion was included in addition to performance sub-dimensions to observe the differential validity of mega and sub dimensions of performance in predicting 'total performance'. (Ghiselli, 1956; Guion, 1965; Dunnette, 1963; Campbell, 1991, 1994). Schmidt and Kaplan (1971) warn that where dimensions of performance show low, zero or negative correlations with each other, use of a composite measure in isolation will mask

the varying determinants of performance (this argument has recently been consolidated by Murphy and Shiarella, 1997). Steiner, Rain, and Smalley (1993) advocate the use of a performance distribution rating rather than a single evaluation of performance. The present study used an overall, global impression of performance as well as sub-ratings for various aspects of performance. Steiner et al argue that the combination of multiple ratings into a single rating may result in distorted representations of performance, and fails to consider differences between typical and maximum performance levels and abilities. The utilisation of global and multidimensional sub-components of performance should generate an impression of both typical and maximum performance for each rater.

A sample of 35 supervisors of the managerial sample provided ratings on the manager's performance. The supervisor / rater characteristics are summarised in table 7.3

Table 7.3 Supervisor (rater) characteristics

Gender	Age	Ethnicity
Male (95% of ratings)	12% aged 30-39yrs	98% white
Female (3% of ratings)	39% aged 40-49	
2% missing data	44% aged over 50	2% missing
	4% missing	

A total of 33 of the sample of 35 supervisors provided overall performance ratings for 135 managers. The largest number of managers rated by a single supervisor was 16 or 12% of the sample. Performance ratings were collected as follows:

Table 7.4 Performance rating information

	Category	Rated
	Overall rating	1-6
Work with people (%)	General performance	%
	Representation of the organisation	
	Supervision of others	
	Organise effective teams	1 – 4
	Communicate ideas effectively	
Work with data (%)	General performance	%
	Work-related knowledge	
	Problem solving ability	
	Planning and organisational ability	1 – 4
Work with objects (%)	General performance	%
	Proficient use of equipment	
	Knowledge of equipment used by peers / subordinates	
	Resource allocation	1 – 4
	Aware of technological innovations	
General motivation and ability	Dedication to organisation	
	Conscientiousness	
	Energy	1 – 4
	Promotion potential	

The present research necessitated the design of a measure that can incorporate performance appraisal across a number of jobs, levels, and industries in order to provide generalisable, comparable results for different types and levels of manager. It is proposed that the application of Fine's (1955) Functional Job Analysis may be one way that job differences can be accounted for. The performance dimensions included are largely consistent with the research into the nature of managerial work, for example, the 'classical management' functions of planning, organizing, deciding, controlling (Fayol, 1949; Koontz, 1971). Task-list approaches to management, such as those of Kotter (1982) and Mintzberg (1975) offer insight into the ways in which managers spend their time but offer limited scope for understanding the rationale and the motivations beneath the overt behaviour.

The performance instrument under current review is applicable to all levels and functions of management. This instrument is dynamic and may naturally incorporate changes in styles of management, for example, the move towards 'Human Resource' models of management (Cascio, 1995), which may influence and change the actual nature or outcomes of management in relation to 'people', 'things' and 'data' without any resulting threat to validity or reliability.

7.4 Hypotheses

The central hypothesis of this work is that construct-oriented biodata analogue models of psychometric constructs should show incremental validity over traditional psychometric measures of these constructs. The concurrent validity study of Watson Glaser, extroversion and neuroticism scores and analogue models of these measures will assess the predictions outlined above. This procedure will involve the calculation of criterion-related validity of both the biodata analogue models of the psychometric constructs and the predictive validity of the traditional psychometric measures. Total performance rating and a career progress variable (number of promotions / time in organisation *100) will be used as criterion measures.

This prediction can be broken down to observe the component parts of this thesis:

1. ***The biodata analogue model of critical thinking ability will show incremental validity over the Watson Glaser Critical Thinking Appraisal test measure of this construct.***

1.A It is hypothesised that critical thinking ability, as assessed by the Watson Glaser Appraisal and the biodata analogue, will correlate positively and significantly with career progress and with overall job performance. It is likely that this construct should correlate significantly with the subcomponents of work with 'data' and work with 'things'.

2. ***The biodata analogue of extroversion will show incremental validity over the Eysenck Personality Inventory measure of this construct.***

2.A It is hypothesised that extroversion, as assessed by the Eysenck personality Inventory and the biodata analogue models, will correlate positively and significantly with career progress and with overall job performance rating. It is likely that this construct should correlate significantly with the subcomponents of work with 'people'.

3. ***The biodata analogue of neuroticism will show incremental validity over the Eysenck Personality Inventory measure of this construct.***

3.A It is hypothesised that neuroticism, as assessed by the Eysenck Personality Inventory and the biodata analogue models, will correlate negatively and significantly with career progress and with overall job performance rating. It is likely that this construct should correlate significantly with the subcomponents of work with 'people', 'data' and 'things'.

Chapter 8 Analyses and results

This chapter presents the results of the primary focus of the research. In the interests of coherence several approaches to ascertaining the validity of the biodata analogue models are discussed (after Cronbach & Meehl, 1955).

Biodata analogue scale reliability statistics are then presented in section 8.2, followed by analysis of construct validity to determine if the biodata analogue models predict the same constructs as the psychometric tests. Concurrent criterion-related validity is then assessed to compare the performance of both the original psychometric tests and the biodata analogue approach to prediction of performance. The incremental validity of the biodata analogue models is then considered and the hypotheses are directly addressed. The validity of analogue models designed purely for the prediction of total performance ratings and career progress ratings (a traditional empirical approach to biodata development) is reviewed in chapter 9. This section is followed by a brief summary of the results.

8.1 Pilot study results

All items with less than 15% of the sample responding were excluded from the scales in an attempt to maximise scale reliability. Table 8.1 presents the alpha coefficients and split half reliability results for Mitchell's scales.

Table 8.1 Reliability analyses on rational biodata models developed from the Mitchell sample of 'Potential Managers'

Scale	N	No. of items	Alpha coefficient	Standardised item alpha	Test-retest correlation for scale
CTA	227	11	0.436	0.537	0.957 (11) P=0.000
Extroversion	223	28	0.589	0.605	0.925 (9) P=0.000
Neuroticism	227	17	0.621	0.623	0.859 (11) P= 0.001

The final column of table 8.1 shows the test-retest reliability of the questionnaire was determined by inviting a sub-sample (n = 12) of participants from the initial study to complete the biodata questionnaire after a period of approximately 12 months.

Table 8.2 Rational biodata analogue scale test-retest reliability

Scale	Alpha Coefficient	Number of items
Critical Thinking Scale	0.47	19
Extroversion	0.65	27
Neuroticism	0.62	21

The alpha coefficients obtained in this set of biodata analogue models are similar to those from the models developed on the pilot sample in this study. This finding offers additional supportive evidence that construct-oriented biodata analogues show a moderate level of internal consistency approaching that of standardised psychometric tests. (Please refer to section 8.1 for a discussion of biodata alpha levels and pilot study results).

Table 8.2 shows the results of the correlation between the weighted-item scale prediction of each psychometric construct collected at time#1 and the rational model prediction collected by completion of the biodata form at time#2. This correlation represents results for 5% of the original pilot sample.

The scales may be said to demonstrate reasonable levels of reliability. The alpha levels are approaching comparability with the usual criteria of 0.8 as an acceptable level of reliability for questionnaires (after Cronbach, 1951). Although none of the three scales meet the suggested level of 0.8 by Hough and Paullin (1994), biodata analogue scales would not usually be subjected to internal reliability analysis and Mumford and Owens (1987) suggest that the alpha levels of biodata scales might be expected to be low as the items are usually developed in an attempt to capture a wide range of variance rather than a single psychological construct as is more often the case with standard psychometric test measures.

8.1.1 Validity of the pilot study models

In addition to the cross validation exercise new biodata analogues of the psychometric constructs were developed using the same techniques (based upon correlation analysis) as Wilkinson (1995). The analogue models predict the psychometric test scores as illustrated in table 8.3

Table 8.3 Construct validity of the biodata analogues developed on a sample of 'potential managers'

Construct	Mitchell	
	Weighted item	Unit weighted
Critical Thinking ability	0.418 (159) P=0.000	0.403 (159) P=0.000
Extroversion	0.711 (222) P=0.000	0.683 (222) P=0.000
Neuroticism	0.559 (226) P=0.000	0.521 (226) P=0.000

The scales developed to predict extroversion appear to predict the psychometric measure of this construct best. This scale also has the most items in (see Table 8.3 above) but is not the most reliable.

Wilkinson (1995) reports correlations between the analogues and the psychometric tests between 0.56 - 0.63 for the Watson Glaser Critical Thinking Appraisal, and 0.64 - 0.68 for the Eysenck Personality Inventory measure of Extroversion (Wilkinson, 1995). Several possible explanations may underlie the differences in the validity between the current model of critical thinking ability and that reported by Wilkinson (1995). This may be due to several factors such as differences in educational attainment between Wilkinson's sample and the current sample such as age of the samples, place of pre-university education, sample size, and the number of items in both models. Further exploration in the cross validation exercise will illuminate if Wilkinson's model is a better predictor of critical thinking ability on the present sample than the current model, thus highlighting potential flaws in the current model (see Table 4.7). A table showing common items in the models may also highlight reasons behind this disparity (see Table 8.5).

Wilkinson's critical thinking analogue was developed on a sample of 149 'potential managers' and consists of 20 items in comparison to the current 11 items. Wilkinson's sample comprised of 55% males and 45% females from a sample in comparison to the current 42 % male and 58% female. In relation to age distribution, 35% of Wilkinson's sample were under 20 years, 49% between 20-30 years, 9.6% 30-39 years, and 10% were over 40 years. Though Wilkinson does not report the percentage of participants completing their pre-university education outside of the UK, the samples do not appear to differ significantly in terms of demographic characteristics and the mean scores on the psychometric instruments are very similar; the mean critical thinking score for Wilkinson's sample was 57.1 (standard deviation 8.4) in comparison with a mean of 56 and a standard deviation of 8.19 on the current sample.

The critical thinking model developed by Mitchell incorporates 11 items, seven of which

were present on Wilkinson's original biodata questionnaire. Mitchell's extroversion model incorporates 28 items, six of which were present on the original questionnaire, the majority being added in the extended version used for this research.

Table 8.4 presents the cross validation coefficients from the application of models developed by Wilkinson (1995) onto the current sample of potential managers.

Table 8.4 Cross validation of Wilkinson (1995)

Construct	Weighted item	Rational
Critical Thinking ability	0.243 (114) P=0.009	0.214 (114) P=0.022
Extroversion	0.504 (232) P=0.000	0.451 (232) P=0.000

It can be observed that both the critical thinking and extroversion models appear to have suffered shrinkage. The original correlations reported by Wilkinson are 0.56 weighted-item for WGCTA using 20 items, 18 of which were transferable onto the current sample; and 0.68 for Extroversion weighted item using 17 items, 16 of which were transferable to the current analysis.

Table 8.5 illustrates which of the items appear in models developed by both Wilkinson and Mitchell, and the direction of the correlation with the psychometric construct. NB [To appear in the biodata models, the item must demonstrate a significant correlation (P=<0.05) with the psychometric measure of the construct].

The items present in table 8.5 fit with the intuitive conception of each of the psychometric constructs under investigation. This table demonstrates that the biodata analogue scales may be likely to have face validity.

relatively good performance of Wilkinson's model of critical thinking, developed on a similar sample of 'potential managers'. This suggests that the samples differ in terms of their critical thinking. This may reflect a difference in the numbers of participants who did not speak English as their first language in one or other of the samples.

Table 8.6 demonstrates the cross-validity of biodata analogue models developed on two different samples of 'potential manager's' for the prediction of the psychometric test scores of 'practising managers'. It can be noted that all models give significant predictions. However, the critical thinking analogue of Wilkinson (1995) is superior to that developed on the present pilot sample while by contrast the extroversion biodata scale is a better predictor that Wilkinson's earlier model. As a result of this it may be concluded that the inclusion of more items for the prediction of extroversion and critical thinking ability improves the predictive power of that analogue. This inclusion of additional items is done rationally in the first instance in the generation of items likely to capture variance associated with critical thinking ability or extroversion respectively, then empirically through the inclusion of only those items demonstrating significant correlations with the test score in the biodata analogues. Therefore the relatively poor performance of the current model of critical thinking ability does not suggest that this construct cannot be captured by biodata following further item generation and model development. The analogue models of neuroticism show reasonable and significant prediction of neuroticism scores for the sample of managers providing evidence of cross-validation of the analogue developed to capture this construct.

Table 8.6 Cross validation of models developed from 'potential manager' samples (Mitchell pilot study, & Wilkinson, 1995) upon a sample of managerial job incumbents

Construct	Wilkinson (1995)			Mitchell		
	Wilkinson original W I models	Weighted item model	Rational model	Mitchell original W I models	Weighted item model	Rational model
Critical thinking ability	0.56 (149) P=0.000	0.45 (116) P=0.000	0.46 (116) P=0.000	0.42 (159) P = 0.000	0.28 (124) P=0.002	0.27 (124) P=0.002
Extroversion	0.68 (149) P=0.000	0.57 (132) P=0.000	0.51 (132) P=0.000	0.71 (222) P = 0.000	0.69 (130) P=0.000	0.63 (130) P=0.000
Neuroticism	-	-	-	0.56 (226) P =0.000	0.45 (134) P=0.000	0.40 (134) P=0.000

All of the analogue models show some shrinkage in validity. Though in general this is lower for the weighted item models and this finding lends some support for the use of the weighted item model for development over the unit-weighted approach. The inclusion of actual correlation coefficient weights in the model appears to make this model more robust.

Cross validation analyses support the potential generalisability of validity and reliability from construct-oriented biodata analogue models of psychometric constructs. The cross validity of analogues developed on two developmental samples demonstrate significant correlations with the psychometric test scores from the current managerial sample (see table 8.7). There is further support for the earlier conclusion that critical thinking and neuroticism do not significantly predict either career progress or total performance ratings (see table 8.11) though a cross validation of Wilkinson's model of extroversion suggests that extroversion may be significantly related to managerial performance as indicated by career progress.

Table 8.7 Comparison of levels of psychometric test scores of critical thinking ability, extroversion and neuroticism across three samples of 'managerial types'

Sample	Critical thinking ability	Extroversion	Neuroticism
Wilkinson (1995)	Mean =57.1 St. Dev=8.4 N=167	Mean =13.7 St. Dev=4.5 N=161	uncollected
Mitchell Potential 'managers'	Mean =56 St. Dev=8.2 N=172	Mean =13.9 St. Dev=4 N=242	Mean =10.5 St. Dev=4.7 N=242
Mitchell Managerial job incumbent	Mean =59.8 St. Dev=8.56 N=135	Mean =12.24 St. Dev=4 N=135	Mean =6.7 St. Dev=4.5 N=135

With the exception of the Mitchell model of critical thinking ability, the rest of the correlations (ranging from 0.4 to 0.69) suggest that there may be similar levels of extroversion and critical thinking ability among the three separate samples of managerial applicants and managerial job incumbents. In fact, the mean scores support this notion

Table 8.8 Items developed for the prediction of neuroticism

Item	Neuroticism	Extroversion
Do you worry about mistakes you have made?	**-0.480 (240) P = 0.000**	0.047 (240 0 P = 0.464
Most of the time I am bursting with energy	**-0.130 (240) P = 0.043**	**0.378 (240) P = 0.000**
How important is job security to you?	**-0.232 (239) P = 0.000**	**-0.205 (239) P = 0.001**
Given the opportunity, would you prefer to work full/part-time/not at all	-0.016(240) P=0.800	-0.035 (240) P = 0.585
Are you ambitious?	-0.018 (240) P = 0.780	**-.296 (240) P=0.000**
Is it important that you get recognition for success in your field?	-0.073 (240) P = 0.257	**-.200 (240) P=0.002**
Do you consider yourself to be more of a risk taker than most?	**0.157 (139) P = 0.015**	**-0.338 (239) P=0.000**

It is interesting to observe the efficacy of the items developed to predict neuroticism. In many cases, these items offer better prediction of Extroversion score rather than Neuroticism score. Table 8.8 illustrates the item correlations with the scores of psychometric constructs.

8.2 Central investigation results

8.2.1 Managerial sample descriptors and inferential statistics by gender and age

Male and female managerial job incumbents did not differ significantly in number of promotions, absence, or in length of time at the organisation. There were no significant gender differences for scores of neuroticism, extroversion and critical thinking ability. Similarly, no gender differences emerged in relation to the importance of work with people or data. Interestingly, female managers were significantly more likely to achieve higher overall scores on ratings of job performance relating to 'work with 'things'' by their manager.

The male managers were significantly older than the females and had significantly higher salaries. Female managers received significantly higher ratings on overall job performance, how well they meet the requirements in their job to work with people, ratings for perceived conscientiousness, enthusiasm, energy, and promotion potential.

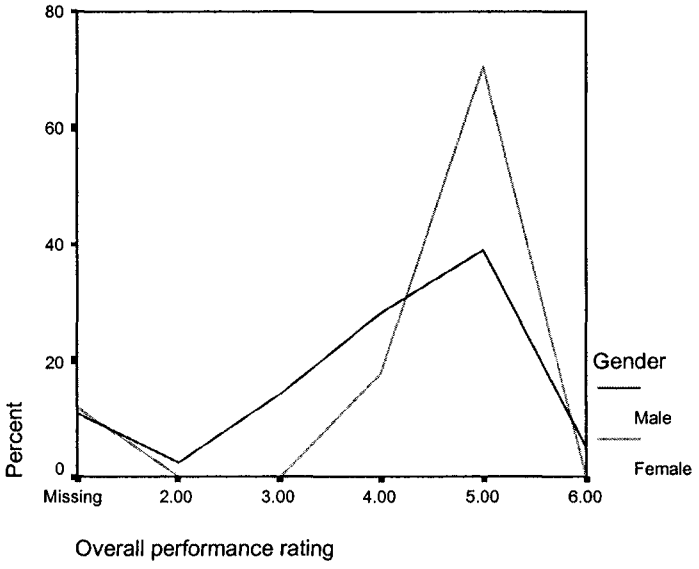


Figure 8.1 Rating of overall performance by gender

Table 8.9 Exploration of gender differences and performance outcomes

Item	Means		df	t	P Level
	Male (n=118)	Female (n=17)			
Neuroticism score	6.7	6.9	20.7	-0.13	0.89
Extroversion score	12.2	12.6	25	-0.45	0.65
Critical thinking ability	60.1	57.7	28.8	1.50	0.14
Age	42.8	34.7	19.39	**3.53**	**0.002**
Salary (£)	23k	14.5k	56.7	**7.86**	**0.000**
Career progress	14.7	14.3	23.2	0.14	0.883
Overall performance rating	4.3	4.8	38	**-3.34**	**0.002**
Work with people (Importance / 100)	47.3	42.5	26.4	1.31	0.198
Work with data (Importance / 100)	37.1	38.7	21.8	-0.42	0.674
Work with things (Importance / 100)	15.6	18.7	21.6	-0.89	0.379
Performance with People (%)	73.0	89.0	30.5	**-6.42**	**0.000**
Performance with data (%)	74.4	83.3	24.4	-1.85	0.076
Performance with things (%)	74.5	77.5	18.1	-0.52	0.606
Perceived dedication	3.3	3.4	21.7	-0.20	0.844
Perceived conscientiousness	3.5	3.8	28.1	**-2.82**	**0.009**
Perceived enthusiasm	3.15	3.59	26.3	**-3.11**	**0.004**
Perceived energy	3.07	3.47	24.4	**-2.33**	**0.028**
Perceived promotion potential	2.88	3.53	32.9	**-4.32**	**0.000**

(N.B. Equal variances have not been assumed. All significance levels refer to a two-tailed t-test)

Figures 8.1 - 8.5 illustrate the variation between male and female managers and the ratings for overall performance and perceived conscientiousness, energy, enthusiasm and promotion potential.

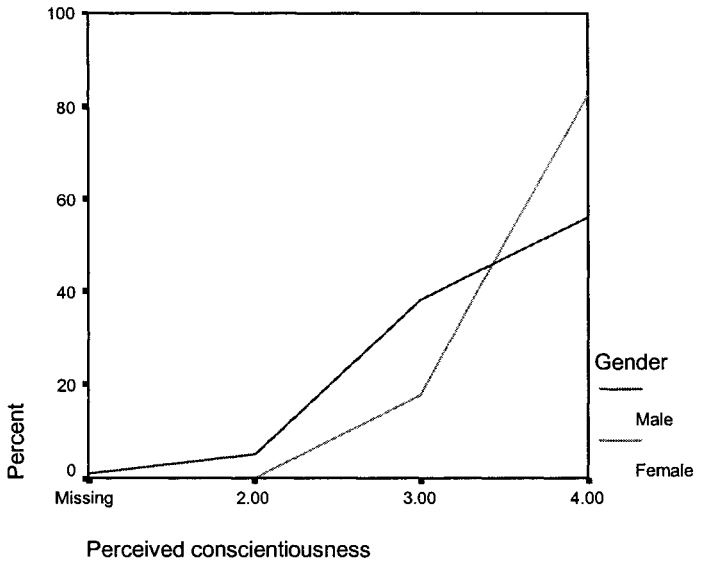


Figure 8.2 Rating of perceived conscientiousness by gender

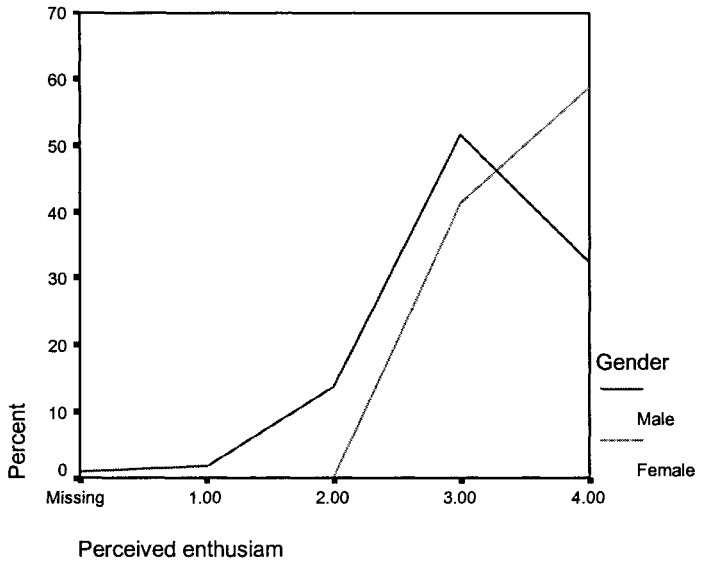


Figure 8.3 Rating of perceived enthusiasm by gender

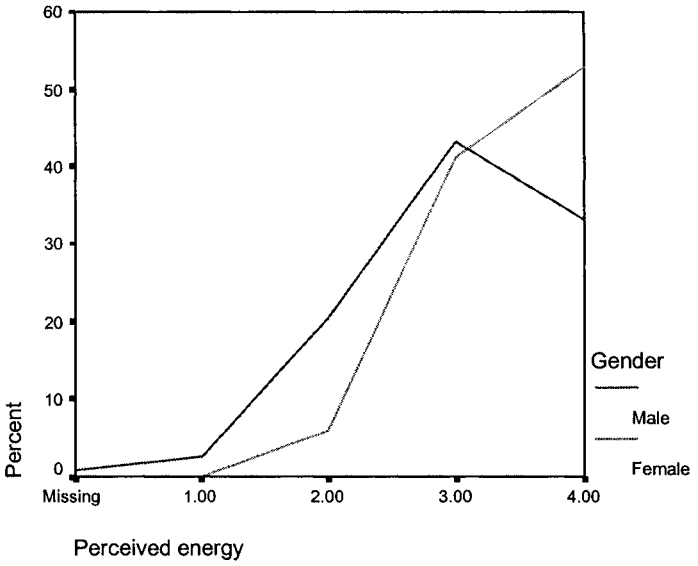


Figure 8.4 Rating of perceived energy by gender

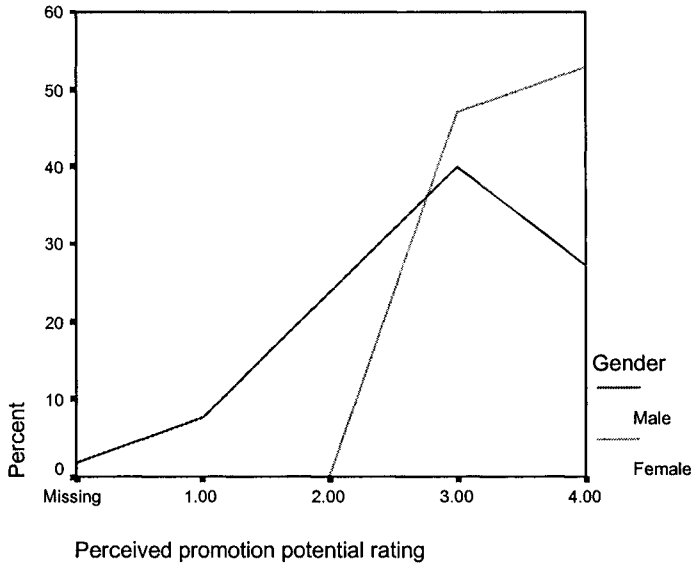


Figure 8.5 Rating of perceived promotion potential by gender

The graphs illustrate the concentration of the ratings of the female managers towards the higher end of each rating scale. In comparison, the ratings of the males appear to have a wider range. This is not necessarily the result of halo error in the ratings of females that can be attributed to their being female for example. The much larger number of males in the samples make it reasonable to suggest that increasing the number of females would be

likely to balance out this positive skew or leniency effect. It is also possible that female managers are perceived as more conscientious, energetic, enthusiastic and with greater potential than their male counterparts and would need to 'stand apart' from the crowd in an organisation dominated by middle-aged males and a low turnover rate for staff.

Correlation analysis revealed that there were no significant correlations between supervisor age or gender and total performance rating or career progress of individual ratees. However, the correlation between gender of the supervisor and rating of 'How dedicated to the organisation is the candidate?' is approaching significance at 0.165 (132) P= .059). This finding suggests that female supervisors perceived their subordinates to be more dedicated than male supervisors. No significant effect was found for either female-male or female-female supervisor-manager relationship. Supervisor gender also correlates positively with number of promotions at r = 0.164 (131) P= 0.061 and is approaching statistical significance though the proportion of female supervisors (3% of all raters) suggests that these findings should be interpreted with caution.

The correlation between supervisor age and number of promotions experienced by the manager is significant at -0.199 (128) P=0.025. This finding may be explained in a number of ways, i.e. younger supervisors may offer more promotions to their managerial subordinates, younger supervisors may be more dynamic than their older counterparts (having attained senior managerial status at a younger age) and therefore their subordinates may enjoy enhanced success and promotion opportunities. The significance of this correlation actually increased when length of time the manager has worked in the organization, and the gender of both supervisor and manager are partialled out (r = -0.227 (123) P=0.011 to control for supervisor age and gender and manager gender).

N.B. [Please note that the term 'supervisor' refers to the person providing performance ratings for the managerial sample.]

8.2.2 Construct validity of biodata analogue models

To begin the incremental validity investigation, it was necessary to build biodata analogue models of the constructs. This was performed in exactly the same way that pilot models were developed for the sample of 'potential managers'. The alpha coefficient was calculated for each scale. Scale items are presented in VII.

In order to investigate the construct validity of the biodata models, it is useful to present correlations of the analogue models with the original psychometric test scores.

Table 8.10 Correlations between analogue models developed from the managerial sample and psychometric measures of the constructs

Construct	Weighted item	Unit weighted / Rational
Critical Thinking Ability	0.219 (113) P=0.020	0.202 (113) P=0.032
Extroversion	0.732 (128) P=0.000	0.704 (128) P=0.000
Neuroticism	0.630 (132) P=0.000	0.617 (132) P=0.000

In general, items found to capture the psychometric constructs fit intuitively with the conception of such constructs. Individual items that were used to build the biodata analogue models are outlined in Appendix VII. There were no significant correlations reported between the original construct measures of extroversion, neuroticism with critical thinking score or gender supporting Eysenck's and Eysenck's (1963) that extroversion and neuroticism are distinct from cognitive ability. However, extroversion and neuroticism correlated significantly and negatively for the managerial sample (r = -0.201 (135) P =

0.019). Similarly, the analogue models for critical thinking ability did not show significant correlations with the analogues of extroversion or neuroticism but those developed on the personality constructs had a significant and negative correlation (r = -0.744 (124) P = 0.000). This finding is in opposition to Eysenck's hypothesis that Extroversion and Neuroticism are orthogonal constructs.

Weighted item analogue models demonstrate stronger correlations with the psychometric test scores than the 'rational' or unit-weighted models. The analogue models of extroversion offer the best of the prediction of psychometric scores. This finding may be due to the extroversion model having more items in it than models predicting critical thinking or neuroticism scores and would then highlight the possibility that these predictions may be improved if further items designed to tap these constructs were added to the biodata questionnaire. It is interesting to note that the extroversion model has the highest internal reliability of the three analogues. Figures 8.6, 8.7 and 8.8 summarise the characteristics captured by each set of models.

Negative correlations	**Positive correlations**
Work with data	Work with People
Have not completed an HE qualification	Have studied statistics
Rate problem solving as weak	Practice until perfect when acquiring new skills
Rate self as a private person	Rate self as a very tidy person
Need supervision when undertaking a new task	Member of a professional or learned society
Rate job security as relatively unimportant	Reading is the best way for you to
You experiment with controls with new equipment	learn about a subject of interest

Figure 8.6 Items included in the critical thinking model

Negative correlations	Positive correlations
Have never considered setting up own business	Compete regularly in sport
More a listener than a talker	Sports or club management in last 3 years
More a thinker than a doer	Describe self as a very sociable person
In a disagreement I tend to keep quiet	Compared to others, I am very carefree
I am comfortable as I am (rather than ambitious)	Most of the time I am bursting with energy
I am not a risk taker	Job security is unimportant to me
It is not important for me to get success & recognition for success	Group work is an effective for me to learn
	I explore a range of approaches in my work
I read about new subjects on interest	I work in short spells on demanding tasks
I am practical thinker	In general I am an untidy person
Work best as an individual	I prefer the role of 'driving force' in teams
I practice new skills until adequate	I am continuing education to make best use of my skills
Increasing age	
Tackle problems in logical stages	
Prefer to work with numbers or symbols	

Figure 8.7 Items included in the extroversion model

Negative correlations	Positive correlations
I worry about mistakes I make occasionally	I work best as an individual
I need minimal supervision on new activities	I rate my artistic ability as weak
I prefer unpredictable project work	I never attempt to repair broken mechanical devices
Job security is unimportant to me	
Compared to others I am carefree	
I can work on many tasks at once	
Increasing age	
Completed a course of study at a Polytechnic	
Most of the time I am bursting with energy	
When things get difficult I try to find a new way around the problem	

Figure 8.8 Items included in the neuroticism model

8.3 Concurrent validity of biodata analogues and psychometric tests scores when predicting managerial performance

A multidimensional approach was adopted with the collection of both subjective supervisor ratings of performance and personnel data that was combined to produce a 'career progress' variable (number of promotions divided by length of time in the organisation multiplied by 100). The concurrent validity of the analogue models and the original psychometric measures was examined by correlating each performance index with the original psychometric measures of critical thinking ability, extroversion and

neuroticism, and with biodata analogues predicting these constructs Table 8.11 presents these correlations. (Appendix VI outlines the performance appraisal proforma).

Table 8.11 The concurrent validity of models developed on managerial job incumbent sample

Construct	Total Performance	Career Progress
Critical Thinking Ability	0.099 (120) P=0.283	-0.043 (134) P=0.621
Weighted item model	-0.128 (100) P=0.204	-0.173 (112) P=0.068
Unit weighted model	-0.131 (100) P=0.194	-0.173 (112) P=0.068
Extroversion	-0.011 (120) P=0.902	0.038 (134) P=0.664
Weighted item model	0.023 (113) P=0.811	0.152 (127) P=0.087
Unit weighted model	0.043 (113) P=0.654	0.156 (127) P=0.080
Neuroticism	-0.069 (120) P=0.454	-0.039 (134) P= 0.656
Weighted item model	-0.086 (117) P=0.354	-0.100 (131) P=0.258
Unit weighted model	-0.073 (117) P=0.432	-0.080 (131) P=0.361

*'CP' = number of promotions / length of time in the organisation * 100*

As illustrated in table 8.11, none of the biodata analogue models or the original psychometric test scores significantly predicts either supervisor ratings of performance or the career progress variable.

The raw data illustrates the correlations of each psychometric construct with all performance outcomes and this exercise reveals partial support for the hypotheses regarding the role of individual differences in critical thinking ability, extroversion and neuroticism as predictors of managerial performance.

It is interesting to note that the original measure of critical thinking ability correlates positively with performance rating and negatively with career progress, both the weighted item and rational biodata models predicting critical thinking ability correlate negatively with both performance outcomes. This finding is contrary to the original prediction that

critical thinking would be positively associated with managerial performance. It is interesting to note a significant negative correlation between Watson Glaser score and the supervisor's perception of the importance of work with people when work priorities between work with people, data and things are split to form 100%. The results also reveal a positive correlation with supervisor ratings of the importance of work with data (0.206 (134) P = 0.017) and with salary (0.22 (134) P = 0.01)

The reverse pattern is true of the correlations between indicators of extroversion. Here the original measure correlates negatively with performance rating but positively with career progress and the analogue models of extroversion correlate positively with both performance measures. Although extroversion correlated negatively though insignificantly with total performance, the pattern of the results offers some support for the prediction that extroversion is likely to be positively correlated with managerial performance. For example, there is a positive significant correlation with supervisor perception of work with people as a job priority and extroversion score (0.203 (134) P = 0.019). This finding suggests that managers with an apparently (relatively) high level of extroversion are likely to be recruited and promoted into roles that involve a relatively high proportion of prioritised work with people.

Further analysis of the neuroticism correlation matrix revealed some supportive evidence for a negative relationship between N score and performance though neither total performance rating or career progress was significantly related to N score. In relation to work with 'people', neuroticism had a significant negative correlation with the item 'Can this person organise effective teams that get the job done?' (-0.179 (128) P = 0.043). With regard to work with 'things', items 'Is this candidate proficient in use of equipment

required for the job?'(-0.243 (131) P = 0.005) and 'Does this candidate demonstrate knowledge of equipment used by subordinates / peers?' (-0.268 (127) P = 0.002) correlate negatively and significantly. The general motivation questions relating to energy and enthusiasm also suggest a significant negative correlation between supervisor perception of motivation and employee neuroticism score. 'How much enthusiasm does this person show consistently?' (-0.179 (134) P = 0.038) and 'How much energy does this person consistently show?' (-0.218 (134) P = 0.012)

It is interesting to note that in each case (with the exception of critical thinking score and performance ratings and career progress), prediction of career progress achieves larger correlations and higher significance levels. Students' t-test and correlation analyses reveal several significant results by rater gender and age (see section 8.2), one example is the possibility that younger supervisors may be more likely to promote their (managerial) subordinates than older supervisors. It is likely that the career progress, providing a standardised criterion for all participants, may reflect truer differences in participant performance relating to individual differences rather than capturing the possible effects of interpersonal processes such as liking and perceived similarity that may confound supervisor ratings.

In each instance, the biodata analogue models offer larger correlations with improved probability levels than the original test scores. These findings offer hope for the primary hypothesis:

Biodata analogue models demonstrate incremental validity over original psychometric test measures.

8.4 Incremental validity of biodata analogue models

The incremental validity of the biodata models will be ascertained by performing a linear regression on the psychometric constructs and the biodata analogue models predicting these scores. The regression data is presented in tables 8.12, 8.13, and 8.14.

Incremental validity was calculated using the Schmidt and Hunter (1998) formula whereby the original *r* levels are deducted from the *r* achieved by combining the analogue model with the original test score, the difference is then divided by the original *r* and multiplied by 100 to attain a percentage increase/decrease. For example, for the incremental validity of the critical thinking analogues, the incremental validity coefficient is the difference between the multiple *r*-values for the original critical thinking score (CTA) and the CTA score in combination with either the weighted item or unit-weighted analogues. The % increase column indicates the percentage of increase in validity to be gained from the addition of the biodata analogues. Dividing the gain in validity figure by the original validity and multiplying by 100 produces this figure.

Table 8.12 Regression data for the incremental validity study on the biodata analogue models of critical thinking ability

	Construct	Original Validity r	Multiple validity r	Gain in validity	% Increase
Total Performance	Critical Thinking Ability	0.099			
	Weighted Item Model	0.099	0.190	0.091	91
	Unit weighted model	0.099	0.190	0.091	91
Career Progress	Critical Thinking Ability	0.043			
	Weighted Item Model	0.043	0.173	0.13	302
	Unit weighted model	0.043	0.173	0.13	302

Table 8.12 outlines the incremental validity analyses for the prediction of both total performance rating and career progress ratings. Both the unit weighted and the weighted item analogue models demonstrate incremental validity over the original Watson Glaser measure of critical thinking for the prediction of Total Performance rating, the significance level of both models being higher than that of the original test score. Similarly, but to much greater effect, both models demonstrate significant increases in validity in comparison to original critical thinking score for prediction of the career progress variable although the starting point for prediction is very low for both the psychometric test measure and the analogue models of critical thinking. It is interesting to note that none of the regression coefficients are significantly predictive of either performance outcome, in contrast to the predictions outlined in the introduction.

Table 8.13 considers the efficacy of the biodata analogue models developed for the prediction of Extroversion.

Table 8.13 Regression data for the incremental validity study on the biodata analogue models of Extroversion score (E)

	Construct	Original Validity r	Multiple validity r	Gain in validity	% Increase
Total Performance	Extroversion	0.011			
	Weighted Item Model	0.011	0.057	0.046	418
	Unit weighted model	0.011	0.080	0.069	627
Career Progress	Extroversion	0.038			
	Weighted Item Model	0.038	0.197	0.159	418
	Unit weighted model	0.038	0.195	0.157	413

In relation to extroversion, both sets of analogue models demonstrate incremental validity over original extroversion score in the prediction of total performance rating and career progress. Both unit weighted and weighted item models approach significance as predictors of career progress (see table 8.11, 'P' values in the region of 0.08) and show incremental validity over the EPI measure of extroversion. The unit weighted (rational) model in particular, appears to capture additional variance related to performance ratings in comparison with the psychometric test score.

Table 8.14 Regression data for the incremental validity study on the biodata analogue models of Neuroticism score (N)

	Construct	Original Validity r	Multiple validity r	Gain in validity	% Increase
Total Performance	Neuroticism	0.069			
	Weighted Item Model	0.069	0.088	0.019	27
	Unit weighted model	0.069	0.079	0.010	14
Career Progress	Neuroticism	0.039			
	Weighted Item Model	0.039	0.103	0.064	164
	Unit weighted model	0.039	0.081	0.042	107

As in the prediction of the earlier constructs, both the unit weighted and weighted item biodata analogue models demonstrate incremental validity over the EPI measure of neuroticism. The gain in validity is substantially lower for prediction of performance ratings compared with career progress and the gain in the neuroticism analogue is lower than that achieved by the analogues of extroversion and critical thinking.

8.5 Summary

The results of the pilot study confirm the construct validity of the biodata analogue models, as they predict scores on the psychometric tests rather well. These results compare quite favourably with the correlation coefficients obtained by Wilkinson (1995) although the extroversion model appears to offer better predictions, most likely due to the addition of new biodata items designed to tap neuroticism. These items somewhat surprisingly correlate with extroversion better than with neuroticism. Table 8.7 demonstrates the new items and their correlations.

Although the critical thinking ability model developed on the current sample of 'potential managers' shows substantial short-term shrinkage (Mael, 1994), this preliminary investigation demonstrates the construct validity of biodata analogues of psychometric measures across three different samples. Such analogue models may offer a transportable 'off-the-shelf' selection tool for use in any organization, following a person-specification to ensure that the attributes or characteristics measured are actually related to job performance.

The results obtained to date suggest that the biodata analogues have considerable construct validity (correlations between the psychometric scores and the biodata predictions of these scores are in the region of 0.47 to 0.69) for the prediction of critical thinking ability, extroversion and neuroticism. This study forms the basis of the incremental validity study as it should illuminate any problems and check the comprehensibility of the instrument.

The second stage of the research involved testing the central hypothesis relating to

incremental validity. A summary of these investigations follows below.

In the main part of the study, biodata analogue models were developed for the prediction of critical thinking score (assessed using the Watson Glaser Critical Thinking Appraisal), extroversion and neuroticism (EPI). These models demonstrate significant levels of construct validity. Table 8.10 presents the correlations between each analogue model and the original psychometric measure of the constructs.

A concurrent validity study was performed to observe potential significant relationships between the original psychometric constructs and biodata analogues with performance criteria. Neither the original psychometric measures nor the biodata analogues produced a significant correlation with subjective supervisor ratings of performance or career progress.

The possibility of the biodata analogue models having incremental validity over the original psychometric measure was then explored. It appears that biodata analogue models of critical thinking ability, extroversion and neuroticism are capable of capturing additional variance not measured by the psychometric tests. This finding is apparent in the incremental validity (increasing levels of significance) of predictions of career progress for both weighted item and unit weighted ('rational') models. Overall, none of the psychometric constructs under investigation demonstrated a significant relation to the prediction of total performance ratings or career progress.

Chapter 9 Additional findings

This chapter provides an overview of significant and interesting findings that were ancillary to the original incremental validity study. The results of an investigation into biodata items that correlate significantly with supervisor ratings and career progress are presented and the findings presented in this chapter also inform the debate on the generalisability of biodata validity by commenting on cross validation results from the present sample to Wilkinson's (1995) original biodata analogue models of critical thinking and extroversion. Regression analysis upon total performance rating and the career progress variable informed an attempt to uncover the components of managerial performance.

9.1 The components of supervisor ratings of managerial performance

Figure 9.1 outlines significant inter-correlations between total performance rating and sub-ratings of performance in ascending order of magnitude. 'Perceived promotion potential' correlates most strongly with overall rating, closely followed by the quality of supervision the manager provides for subordinates. The range of significant intercorrelations with sub-ratings of performance, suggest that the halo error is likely to account for some of the variance across total performance ratings.

Positive correlations	
$r = \mathbf{0.791}$	**Perceived promotion potential**
$r = 0.726$	Quality of supervision provided to subordinates
$r = 0.708$	Problem solving ability
$r = 0.661$	Can organize effective teams
$r = \mathbf{0.652}$	**Consistent energy**
$r = \mathbf{0.626}$	**Consistent enthusiasm**
$r = 0.614$	Perceived conscientiousness
$r = 0.613$	Resource allocation
$r = 0.595$	Candidate representation of organisation
$r = 0.587$	Rating of work involving people
$r = 0.518$	Dedication to the organisation
$r = 0.511$	Communicate ideas effectively
$r = 0.575$	Planning and organizational ability
$r = 0.503$	Rating of work involving data
$r = 0.488$	Technical innovation awareness
$r = 0.404$	Rating of work involving things
$r = 0.397$	Knowledge of equipment used by subordinates / peers
$r = 0.348$	Proficient use of equipment
$r = 0.230$	Work-related knowledge

Figure 9.1 Sub-components of performance: evidence of a halo error

Length of time in the organisation, salary and number of actual promotions did not correlate significantly with total performance rating but did have a significant correlation to career progress as one might expect. It is interesting to note the absence of any obvious real common variance between the two types of performance outcome data.

Positive correlations	
$r = 0.774$ (134) $p = 0.000$	Number of promotions
$r = 0.185$ (133) $p = 0.033$	Work-related knowledge
Negative correlations	
$r = -0.203$ (133) $p = 0.019$	Importance of work with data
$r = -0.207$ (128) $p = 0.019$	Supervisor age

Figure 9.2 Sub-components of career progress

As one might expect, number of promotions is the single best correlate of 'career progress'. It is striking that only three other aspects of performance data correlated

significantly with this variable, one of these being supervisor age. According to these results, the older the supervisor providing the total performance rating, the lower the career progress scores. This finding supports the earlier result suggesting that younger supervisors were more likely to rate their younger subordinates (managerial job incumbents) as worthier of promotion than older supervisors (as discussed in chapter 8, the correlation between supervisor age and number of promotions experienced by the manager is significant at r = -0.199 (128) P=0.025).

This table also suggests that the more importance attributed to work with data, the lower the career progress outcome. For some reason, perhaps relating to the personality trait of introversion – extroversion, managers in this kind of work were unable to enjoy the promotion potential available to managers working primarily with people or equipment. It is also likely that these managers may have occupied a support or maintenance management role for the organisation rather than a dynamic or 'frontline' function involving representation of the company. This difference is not significant when analysed by a t-test (equal variances not assumed: t = -1.287 (df = 27.32) p = .209 (two-tailed).

9.2 The criterion-related validity of empirically derived biodata models

9.2.1 Prediction of total performance rating and 'career progress' score

In order to observe which personal characteristics may be significantly related to supervisor ratings of performance, biodata items were subsequently used to build a 'post-hoc' empirical biodata analogue model of total performance. Using the procedure outline in chapter six, rational and 'unit weighted' analogue models were developed.

Empirically derived models of career progress and total performance based purely on biodata item correlations were calculated. Table 9.1 illustrates the correlations between empirical biodata models and original scores for total performance rating and career progress.

Table 9.1 Empirical biodata analogues of total performance and career progress

	Correlation with original (TP and CP)	No. of items	Alpha level
Total performance ('totp_emp') weighted item scale correlated with 'total performance'	0.407 (115) P = 0.000	6	0.19
Career progress ('cp-emp') weighted-item scale correlated with 'CP'	0.216 (131) P = 0.013	6	0.05

Table 9.1 provides evidence of highly significant correlations between empirically derived models of performance outcomes and the actual outcomes, despite both models having few items in and unacceptable levels of alpha coefficient reliability.

9.2.2 Prediction of perceived conscientiousness and energy

Table 9.2 presents the results from empirically derived biodata analogue models of perceived conscientiousness and energy. Observing individual biodata items with significant correlations with conscientiousness and energy formed the basis for the development of these models. As in all other instances of model development, items with response rates of less then 10% were omitted from the models.

Table 9.2 'Post-hoc' development of empirical biodata models predicting apparent conscientiousness and energy

	Correlation with original rating	No. of items	Total performance	Career progress	Alpha level
Analogue of conscientiousness weighted item scale	**0.54 (126) P = 0.000**	13	**0.30 (113) P=0.001**	0.17 (126) P = 0.052	0.49
Analogue of energy weighted-item scale	**0. 25 (128) P = 0.005**	19	.18 (114) P = 0.060	0.16 (.127) P = 0.069	0.49

It can be observed that empirically derived models of perceived' psychological constructs can demonstrate impressive correlation and significance levels with moderate levels of internal consistency. It is possible that biodata item generation based upon psychological constructs, i.e. a construct-oriented approach, would improve these validity figures. A second explanation for the apparent validity and reliability shown 9.2 may relate to the possibility that perceived conscientiousness and energy form part of the traits that the biodata items had originally been designed to capture. It is likely that individual levels of extroversion and emotional stability confound the perception of both 'traits' of conscientiousness and energy. The discreteness of conscientiousness and energy from extroversion and neuroticism is a matter of considerable debate in the current climate and the current trend is for the widespread application of the five-factor model with no real (applied) comparison between five-factor results and those to be obtained from the application of a model offering greater explanatory power of personality, such as that of Eysenck and Eysenck (1963). This contamination of traits and their sub-categories may be particularly likely in the case of autonomous work characterised by relatively low levels of interpersonal contact with line managers and supervisors.

Table 9.3 Items correlating significantly with perceived conscientiousness (In ascending order of magnitude)

Item	*r* value	DF	P value	Interpretation
Q7-5	-0.258	132	0.003	Preference for maths at school
Q24	-0.244	133	0.005	Attempt to solve problems in their entirety
Q2	-0.243	134	0.005	Age
Q30-2	-0.242	134	0.005	Prefer to have some supervision for new activities
Q46	-0.227	134	0.008	Sports club management
Q10c-3	0.222	134	0.010	Have studied HNC / HND
Q16	0.215	134	0.013	Studied for pleasure: evening/part-time
Q30-3	0.204	134	0.018	Need minimal supervision for new activities
Male	-0.178	134	0.039	Gender
Q13-2	0.176	129	0.046	Concentrate on weak areas as a revision strategy
Q14-1	-0.172	134	0.046	Completed a full course of study at university
Q23-4	0.172	134	0.046	Attempt to relate unfamiliar problems to known solutions

Table 9.4 Items correlating significantly with perceived energy (in ascending order of magnitude)

Item	*r* value	DF	P value	Interpretation
Q2	-0.272	134	0.001	Age
Q59	0.243	134	0.005	Worry about mistakes often
Q16	0.226	134	0.009	Studied for pleasure: evening/part-time
Q21-1	0.220	133	0.011	Prefer to work with a small team
Male	-0.206	134	0.017	Gender
Q5a	0.204	134	0.018	Importance of work with people
Q8-5	0.203	134	0.019	Continuing education to make best use of my skills
Q5b	-0.200	134	0.021	Importance of work with data
Q61-1	0.199	134	0.022	I learn most effectively from loosely supervised project work
Q60	0.197	134	0.022	Most of the time I am bursting with energy
Q26-1	0.195	134	0.024	I like to tackle problems in logical stages
Q64	-0.194	134	0.025	I am comfortable as I am (ambitious item)
Q63	-0.193	134	0.026	I would prefer to not to work at all (Full-time – not at all)
Q28-2	-0.188	134	0.029	Practice new skills until adequate
Q3	0.186	131	0.033	Middle child in the family
Q32-4	-0.179	134	0.038	Learn best about new subject by talking to experts
Q24	-0.174	133	0.045	Attempt to solve problems in their entirety
Q52	0.172	134	0.047	Compared to others I am private (vs. sociable)
Q61-4	0.170	133	0.050	I learn most effectively from group work and discussion

Table 9.3 and 9.4 illustrate the biodata items that correlate significantly with perceived conscientiousness and perceived energy. Several of these items appear in the biodata analogue models of extroversion and neuroticism such as ambition (comfortable as I am versus ambitious) and describing oneself as a private (versus sociable) person in the neuroticism model and the empirical analogue of perceived energy. Experience in sports or social club management features in both the extroversion analogue and the conscientiousness model though in opposing directions. (Refer to appendix VII for the biodata analogue models developed on the managerial job incumbent sample).

9.3 Conversations with managerial employees

As part of the exploration into the failure of critical thinking ability, extroversion and neuroticism to predict total performance rating and career progress a small number of semi-structured conversations were conducted in the interests of verifying the conclusions on the predictions of managerial performance rating and career progression. Informal discussions were held with nine of the managerial job incumbents from one organisation. These meetings took place in April 2001 approximately three years and three months after the first round of managerial incumbent data was collected. All managers taking part in this section of the research volunteered to do so and all responses were maintained in strictest confidentiality. A dictaphone was used to record the conversations with the volunteers consent. Brief notes were made during the discussion and the audiotape recording was used to inform the writing up of these notes. Each conversation lasted between twenty and thirty-five minutes and was conducted on an individual basis with the researcher. This work was not intended to form part of a full-scale qualitative exploration

but served as a small check on the realism of the findings and offered the opportunity to provide a further debriefing session to interested participants.

All managers were asked to reflect upon the possibility of a 'prototypical manager' and to delineate the defining characteristics of managers. This suggestions was met with a range of responses; some of the participants offered stereotypical outlines of managers they interacted with on a daily basis, others offered an 'ideal manager' template. They were encouraged to discuss their perceptions of promotion potential and asked to describe the characteristics that they would favour if selecting for a managerial position. There follows a brief summary of the conversations from the audiotape records.

9.3.1 Summary of conversation data

Previous participants (n = 5)

P1. (Male aged 30-35, in the organisation for approximately 6 years)

The prototypical manager is 'in control of the situation, knows when to take charge and when to delegate. (The manager) should know when to encourage and bring the best out in staff', and is likely to be sociable. This manager sees promotion potential being a function of good knowledge of the organisation, good interpersonal skills, and good communication with all levels of employees. If selecting for a manager, this participant would prefer a candidate who is able to integrate and get on well with others, works well as a team player, has excellent communication skills, and shows initiative. The managerial candidate would ideally have good time management skills and be conscientious about work.

P2. (Female, 30 – 40 years, has worked in the organisation for approximately 5 years)

This participant viewed the management structure as defined by continuous change without reflection. The prototypical manager was perceived as being inexperienced and having no knowledge of the role of his/her subordinates. In order to achieve promotion, this participant discussed the role of decisiveness and single-mindedness.

P3. (Male, 50 – 60, 27 years in the organisation)

Prototypical management practice is viewed as essentially reactive and crisis focussed rather than dynamic or strategic. Promotion was perceived as being linked to being 'in the right place at the right time' and said that the hierarchy was rather rigid with little crossing between work functions between, for example, technical and interpersonal skills-based jobs.

P4. (Male aged 50-60years, has worked within the organisation for just under 20 years)

This participant views the prototypical manager as male, lacking interpersonal skills, poor communication skills, 'trying to do their best but not quite succeeding'. Promotion is perceived as being linked to ability, hard work and good luck. This participant would seek a managerial candidate with the ability to mix well with others, deal with variety and diversity in job tasks and people. The primary factors for managerial potential were thought to relate to motivation and flexibility, 'doing the job to the best of their (sic) ability, conscientiousness'. Managerial candidates should also offer patience, self-discipline, technical ability, computer literacy, articulacy and listening skills.

P5. (Male, aged 30 – 40, in the organisation for more than 15 years)

This participant said it was impossible to give a description of the prototypical organisational manager as there were to o many varied styles of management within the company. If selecting for a manager, this participant would prefer a candidate with good technical ability, understanding of information technology and applications of IT in a function appropriate to the type of work required, self-motivation and conscientiousness. Actual promotion in the organisation was said to be related to dedication, 'going out of your way...' being good at the job, offering new ideas and getting involved 'beyond the call of duty'.

New employees: (n = 4)

N1. (Male aged 40 – 45, working within the organisation for less than 4 years)

Prototypical manager is over 30 years old, driven and ambitious when younger but now settled into career, male, white, introverted though approachable. This employee felt that the organisation had a 'staid' culture and was resistant to change. As a result, perception of promotion is linked primarily to networking skills within and outside the organisation. This manager would prefer a candidate showing a critical or questioning stance of the status quo, and a candidate that is well motivated, keen and creative when selecting for a managerial position.

N2. (Male, less than 1 year in the organisation)

Prototype manager was viewed as white, male, over 40 years of age, with poor time management skills, suffering stress, 'hard to get hold of', indecisive and continuously attending meetings. The perceptions of promotion involved

promotion across a range of managerial functions so that flexibility is viewed s an important factor. Entrepreneurial traits and a willingness to take advantage of the organisation in terms of sponsorship for further study and training were also perceived as important characteristics likely to lead to organisational success. If selecting for own job, this manager would favour a candidate with good communication skills, time management skills, ability to work independently, and with relevant technical qualifications and/or experience.

N3. (Female aged 25 – 35, 1.5 years in the organisation)

Prototypical manager is one who is 'supportive, encouraging, allowed to grow within the job'. When asked about promotion potential this participant said that she had no idea what senior management favoured in making promotion decisions. If recruiting for a managerial position, this participant would seek a candidate with versatility, enthusiasm, high energy and a sense of belonging to the organisation to enable good representation of the organisation.

N4. (Male, aged 30 – 40 years, in a management role for less than 12 months)

This manager outlined the prototypical manager as white, male, timid, deferential to the board of directors, attempts to develop a good working relationship with staff, is indecisive and cautious. Promotion was perceived as being related to organisational politics rather than any intellectual ability, technical skill or personality characteristics. If selecting for a management role, this participant would prefer a candidate showing confidence, good decision making ability, stability, reflective, a logical approach to getting the job done and excellent technical ability.

9.3.2 Summary of findings across established and new managerial job incumbents

Several of the managers discuss the role of critical thinking ability, for example, P2 talks of 'decisiveness' and P4 and P5 discuss the importance of technical ability. N4 asserts that 'good decision making ability, and a logical approach to getting the job done and excellent technical ability' are important for managerial candidates. Similarly, N2 argues for evidence of 'relevant technical qualifications and/or experience' among managerial candidates and N1 outlines the importance of a 'critical or questioning stance of the status quo' in a ideal managerial candidate.

In terms of the importance of extroversion, ability to get on well with others was associated with the 'ideal managerial candidate' by a number of volunteers in the follow-up conversations (for example see the summary of conversations with P1 and P4). Confidence (sometimes referred to as 'surgency' in the psychometric literature) was put forward by N4 and this suggestion seems to support the literature suggesting a positive association between extroversion and job performance. The lack of criterion-related validity in this research may be due to the nature of the performance appraisal proforma, which may have led raters to focus upon aspects of work not associated with sociability and dominance.

Although neuroticism was not explicitly mentioned during conversations with managerial workers, 'stability' is considered by N4 to be an important attribute of the 'ideal managerial candidate', this use of the term refers to emotional evenness or absence of instability or neuroticism. This interpretation was reflected back to the participant to ensure that the interpretation was correct.

Conscientiousness was discussed by P1, P4, P5, as being important for the 'ideal managerial candidate'. P5 outlines the importance of 'dedication, "going out of your way", being good at the job, offering new ideas and getting involved "beyond the call of duty".' N3 adds enthusiasm and high energy to the ideal characteristics and P4, P5, N1 cite 'motivation' as fundamental to their perceptions of the ideal manager.

9.4 Cross validation findings

Cross validation necessarily forms an essential component of the validation for all approaches to biodata analogue and psychometric test development. To explore the 'robustness' of this approach to biodata development, cross-validation analyses were carried out between the models developed from the current sample of managerial job incumbents, those involved in the pilot sample (see chapter 6) or 'potential managers', and the 'potential mangers' recruited by Wilkinson (1995) in the pioneering research study for this approach to construct-oriented biodata development. Where validity coefficients are of similar levels across different samples, this may be interpreted as supportive evidence for the construct-oriented approach to biodata item generation and interpretation.

It has been observed that all previous biodata models significantly predict the psychometric test scores of the sample of managerial workers. Therefore there is substantial evidence for the cross-validity of the construct-oriented approach to biodata development. The Wilkinson (1995) models of critical thinking ability offer improved prediction of critical thinking score than the pilot study for this research. This may be explained via consideration of differences in the pilot samples. Wilkinson's sample

consisted of 'two groups, existing managers and likely 'new' applicants' for managerial jobs' (Wilkinson 1995:148). The mean critical thinking score from Wilkinson's sample was 57.1 (standard deviation of 8.4, n = 167) in comparison to the mean score of the pilot study in this research was 56, standard deviation of 8.2) from a sample of 172 'potential managers' (see page 125). Forty eight per cent or 111 of Wilkinson's sample were practising 'senior managers' in comparison to the pilot study in this research where 60.5% of the sample were aged between 20-24 years and for the most part, involved in training and postgraduate education as a future managerial applicant. For this reason, it is likely that Wilkinson's sample may be more similar to the present sample of managerial job incumbents. Although Wilkinson (1995) did not develop an analogue model of neuroticism a post-hoc analysis permits the correlation between the items in his original model and parallel items from the current data.

Table 9.5 Internal consistency of biodata analogue models across 2 samples

Model	Current managerial model reliability	Mitchell (pilot models, standardised alpha)	Wilkinson (1995)
CTA	α =0.47 (19 items)	α=0.54 (11 items)	α=0.48 (18 items)
Extroversion	α=0.65 (27 items)	α=0.61 (28 items)	α=0.33 (15 items)
Neuroticism	α=0.62 (21 items)	α=0.62 (17 items)	α=0.33 (14 items)

The Mitchell neuroticism scale shows the greatest generalisability of reliability, this alpha coefficient is only slightly lower than that of the scale developed from this sample (α = 0.62) though the critical thinking scale and extroversion scales are considerably lower than the managerial incumbent scales (α = 0.47 and 0.65 respectively).

Table 9.6 Prediction of performance criteria by models developed on 'potential managers'

Cross validation of (rational) pilot models onto the managerial sample			
Construct	Model	Total performance	Career progress
Critical thinking	Mitchell (pilot)	-0.14 (97) P=0.17	-0.171 (109) P = 0.08
	Wilkinson (1995)	-0.09 (102) P = 0.37	0.029 (115) P = 0.76
Extrov	Mitchell (pilot)	-0.05 (115) P = 0.62	0.105 (129) P = 0.24
	Wilkinson (1995)	-0.01 (117) P = 0.95	**0.217 (131) p = 0.01**
Neurot	Mitchell (pilot)	-0.01 (119) P = 0.91	0.060 (133) P = 0.50
	Wilkinson (1995)	0.08 (117) P = 0.39	0.058 (131) P = 0.51

The cross-validated models were then assessed for predictive validity. Table 9.6 demonstrates the results of correlation analyses. The majority of the results are similar to those outlined in chapter 8 with the models developed on the current sample. The correlations are very small for the most part though it is interesting to note that critical thinking ability is correlated negatively with total performance rating for both of the pilot models and with career progress in the pilot model developed in the current research. The most interesting finding from this analysis is the significant relationship between the Wilkinson (1995) model of extroversion and the career progress scores of the current managerial sample. Further consideration of the sample characteristics may explain this unexpected finding. Almost half (48%) of the population involved in Wilkinson's sample reported themselves as 'senior managers'. The mean extroversion score of this sample was 13.7 (standard deviation of 4.50) this mean is just slightly lower than that of the pilot sample in the present study (13.9, standard deviation 4) and slightly higher the present managerial sample (12.2, standard deviation of 4).

9.5 Summary

The development of post-hoc empirical biodata analogue models of total performance and career progress and perceived conscientiousness and energy supports the thesis that analogue measures of constructs can offer incremental validity, even where the individual completing the biodata instrument does not provide the construct rating or performance outcome variables. Despite the supportive results, one must bear in mind the absence of significant predictive validity of the psychological constructs selected for examination, further research involving a measure of general intelligence can add weight to this conclusion.

A purely empirical 'post-hoc' strategy has been adopted for the development of biodata analogues of perceived conscientiousness and energy and these models demonstrate incremental validity over the original supervisor rating of total performance and career progress variables, lending further support to the hypothesis that construct-oriented biodata analogues may capture additional variance and improve the prediction of constructs, even if, as in this instance, the biodata items and psychometric trait reports are from self and supervisor ratings.

Exploration of inter-correlations between performance rating items reveal the strong possibility that halo error may confound performance rating data despite the attempt to subcategorise performance. The possibility that the personality traits of extroversion and neuroticism and perceived energy and conscientiousness may be confounded has been discussed and will be debated further in the discussion.

Chapter 10 Discussion

10.1 Summary of results and support for the hypotheses

Construct-oriented biodata analogue models predicting the psychometric constructs of critical thinking ability, extroversion and neuroticism appear to have incremental validity for the prediction of supervisor ratings of managerial performance and a career progress variable associated with number of promotions and salary.

10.1.1 Critical Thinking Ability

The biodata analogue model of critical thinking ability will show incremental validity over the Watson Glaser Critical Thinking Appraisal test measure of this construct.

The prediction that critical thinking ability (and biodata analogue measures of critical thinking) would demonstrate predictive validity for managerial performance via total performance ratings and career progress was not supported. Construct-oriented biodata analogue models of critical thinking ability offer enhanced predictions and greater utility than the original test (the biodata questionnaire is also considerably faster and easier to complete) but overall, critical thinking as captured by either approach does not significantly correlate with total performance or career progress for this sample of managerial workers.

The prediction that critical thinking ability would correlate positively and significantly with career progress and with overall job performance and with subcomponents of work with 'data' and work with 'things' obtained only partial support. Overwhelmingly, critical thinking ability does not predict career success or total performance rating for the current sample of managers. This unexpected finding may be explained with recourse to several points of concern over the sample. The participants in this sample represented all levels of management from trainee level to the associate director. It is possible that the range of critical thinking ability was too wide to produce a simple linear relationship with total performance rating, career progress or work with data and things. The possibility that critical thinking ability may be predictive of success in senior management is one that cannot adequately be tested with the current sample. 7 (9%) of the participants in the managerial sample earned over 40k and 25 (34%) earned over 30k. These numbers are too small to produce meaningful correlations. It is also possible that the critical thinking appraisal test assesses a very specific aspect of general intelligence and that the range of skills utilised by the heterogeneous sample of managers in this study clouds the relationship between critical thinking and the functions of management outlined in Scriven et al's 91994) conceptualisation of 'executive intelligence'. A measure of general intelligence may provide clearer evidence that construct-oriented biodata analogues of ability have criterion-related validity in the prediction of managerial job performance. A further possibility may include the chance that the performance appraisal proforma developed for this sample cues the rater to focus upon aspects of performance and/or observable outcomes that are not directly related (with a linear relationship) to critical thinking ability.

In relation to the biodata analogue of critical thinking ability, both the rational and weighted item biodata models demonstrate incremental validity over the original Watson Glaser measure of this construct. The prediction of ratings of career progress is particularly improved. This enhanced prediction of performance ratings may be explained by recourse to a discussion of the nature of the biodata items in comparison to those in the Watson Glaser Critical Thinking Appraisal. Biodata items associated with critical thinking score capture previous experience (number of qualifications, previous experience with statistics etc.), self-perceptions regarding ability (problem solving, mathematical ability etc.) and work preferences (prefer to tackle work in logical stages, prefer to work as an individual etc). These preferences are certainly likely to produce an outcome in terms of work motivation and career (promotional) progress as a result. The individual may have greater influence over career progress than over supervisor ratings obtained from a single appraisal exercise. This naturally begs the question of what contributes towards supervisor ratings. The single best predictor of supervisor ratings was a supervisor rating (1 - 4 Likert scale response) of individual promotion potential, followed closely by perceived conscientiousness level and rated 'energy'. Table 9.1 in chapter 9 outlines items from the performance appraisal questionnaire that demonstrated significant inter-correlations with total performance ratings.

Based upon the follow-up discussion sessions with the sub-sample of managers wherein each manager was asked to describe their ideas about the 'prototypical manager' and the 'ideal' managerial candidate, a measure or test of technical ability may have provided a better, more criterion-valid predictor of performance rating and career progress than the Watson Glaser Critical thinking appraisal. This approach may have offered a better predictor of performance but ultimately may not have been equally appropriate or

generalisable across the different organisations used in the research.

The construct validity of the critical thinking analogue models was evident in the significant correlations of both analogues with original critical thinking score. However, the analogues of critical thinking demonstrate lower construct validity than those capturing extroversion and neuroticism for both the pilot sample models and the managerial samples. Furthermore, the reliability of the analogue models of critical thinking was relatively low, alpha coefficients did not exceed 0.5. This outcome may be related to restriction of range in the samples, the raw data shows a slightly positive skew of the distribution of Watson Glaser critical thinking scores though the mean score (59.8, standard deviation 8.6) is slightly lower than that for managerial samples published by Watson and Glaser (1980) where the published mean was 63.6, the standard deviation was 8.4, in a sample of 106 (95 male and 11 female) senior British managers aged between 25 and 58 years.

10.1.2 Extroversion

The biodata analogue of extroversion will show incremental validity over the Eysenck Personality Inventory measure of this construct.

Incremental validity is demonstrated from both the weighted item and unit weighted / rational biodata models for the prediction of both career progress and total performance rating. The increase in validity is expressed as a gain between 413% to 627% across the biodata models and predictions. These figures represent an impressive gain but the predictive ability of the original measure of extroversion was minimal to begin with. The

rational biodata analogue predicting total performance ratings offers the largest increase in validity, supporting the primary hypothesis of incremental validity. However, as demonstrated in the critical thinking predictions, none of the extroversion measures predicts either performance outcome significantly. The extroversion models demonstrate impressive construct validity and a good level of internal consistency (observed in the consistent alpha coefficients) suggesting that the analogues do successfully capture extroversion.

The prediction that extroversion would correlate positively and significantly with managerial performance (total performance an career progress) received only indirect and limited support through the supervisor ratings of the importance of work (positive and significant) with people and work with data (a significant negative correlation). However, for the current sample of managerial employees, extroversion does not predict ratings of overall performance or career progress significantly.

10.1.3 Neuroticism

The biodata analogue of neuroticism will show incremental validity over the Eysenck Personality Inventory measure of this construct.

The incremental validity of biodata analogues of neuroticism is clear though the gains (expressed as percentages) are much lower for the prediction of total performance rating (14.5 – 27.5%) in comparison to the analogues of extroversion and critical thinking. The models perform better for prediction of career progress though the gains are still relatively small compared to the prediction of the earlier constructs (107 – 164%).

The construct validity of the analogue models of neuroticism is demonstrated in the significant prediction of EPI 'N' score (see table 8.2) and the items included fit intuitively with the conception of neuroticism or emotional instability (see figure 7.3). The reliability of the rational biodata analogue prediction of 'N' is good (see table 7.2).

The second component of the hypothesis in relation to neuroticism concerned the criterion-related validity of neuroticism and total performance rating and career progress. As predicted, all measures of neuroticism correlated negatively with total performance rating and career progress though none of the correlations reached significance. It is possible in a concurrent validity study such as this that there is likely to be restriction of range in neuroticism scores as the sample consisted of employees that were considered to be relatively successful and had achieved at least junior manager or management trainee status. As a result, employees with higher levels of neuroticism may have been 'selected out' of this sample earlier. The neuroticism data confirms this hypothesis, the mean score was 7 (standard deviation of 4.5), much reduced from Eysenck and Eysenck's (1963) published mean of 11 for managers.

10.2 Construct-oriented biodata

The results support the hypothesis that construct-oriented biodata analogue models should offer additional gains in validity over original psychometric test measures. This finding supports the earlier work of Wilkinson (1993, 1994, and 1995) on a sample of potential managers, and with Mount et al (2000) for a mixed-occupational sample. The current research also supports the trend towards development of construct-oriented biodata analogues for pre-selection (Hough & Paullin, 1994; Mumford et al., 1996).

Appendix VII presents a comprehensive list of all biodata items involved in the analogue models developed for the prediction of critical thinking ability, extroversion, neuroticism, perceived energy and conscientiousness. It is clear that the items relating to self-report of trait like preferences (e.g. q59' I tend to worry about mistakes occasionally- very often') significantly predict a range of the 'personality' variance. It is also interesting to note that items capturing critical thinking ability appear to capture self-reports of ability (e.g. problem solving and mathematical ability as one would expect, but also capture variance associated with more 'trait'-like preferences such as (q48) 'in terms of my general standards, I am a very tidy person – very untidy person'.

The cross-validation study findings suggest that construct-oriented biodata analogues of psychometric constructs may demonstrate generalisability over occupational samples and do not necessarily suffer from excessive shrinkage over time in the way that traditional empirically based biographical history inventories may. Construct-oriented biodata analogues combine the advantages of traditional biographical information, offering samples of behaviour in addition to signs and provide incremental validity compared with psychometric measures of personality constructs. Analogue models developed in this way can also improve understanding of the psychological constructs that may influence performance and contribute to a theory of managerial work as a result.

10.3 The performance criterion

The performance criterion consisted of a multi-trait job analysis approach to attempt to capture performance relating to people, data and things (after Fine & Wiley, 1955). This technique was applied in part, to attempt to avoid 'halo' or 'horns' (Berry & Houston, 1993) errors and attempt to ensure that supervisors providing performance ratings were able to focus on each aspect of work and performance separately.

Overall, the best predictive component of managerial performance as measured by supervisor ratings of overall job performance was perceived promotion potential. In relation to general qualities, perceived energy and conscientiousness and promotional ability were significantly positively correlated with overall performance rating. Overall or total performance rating correlated significantly with the majority of the other sub-ratings suggesting that a halo or horns error is likely to operate in the production and collection of performance ratings, despite an attempt to distil areas of competence into categories relating to work with people, data and things through the application of a Functional Job Analysis model.

The indefinable and esoteric nature of managerial work ensure that the process of drawing up person specifications and applying criteria for appraisal of performance are problematic in terms of obtaining valid indications of what managerial jobs actually involve. In order to develop an instrument for managerial appraisal, four questions were addressed:

1. What is the nature of managerial work?
2. What constitutes effective managerial performance?

3. How can this be assessed?

4. How might individual differences contribute variance to job performance outcomes?

In summary, effective job performance is conceptualised as relating to the situation as well as the individual (after Borman & Motowidlo, 1993, 1997; Findlay, Giles, & Mossholder, 2000; Motowidlo & van Scotter, 1994). This model integrates the importance of individual differences relating to intelligence and personality, motivation and interests but also incorporates the role of interpersonal relationships and political power exchanges, including the performance of organisational citizenship behaviours (Findlay et al., 2000; Konovsky & Organ, 1996; Newman & Kickul, 1998; Organ, 1994, Organ & Konovsky, 1989; Organ & Lingl, 1995; van Dyne, Graham & Dienesch, 1994).

Job knowledge = time at work + intelligence + access to information and important groups for information

Effective performance = job knowledge + intelligence + traits specific to industry or work function

Figure 10.1 Job knowledge and effective performance

It is worthwhile to consider the 'mixed model' of performance offered by Hartle (1995) at this juncture. If job performance is conceptualised as work behaviour and work outcomes, then the stance adopted in this research, in the attempt to predict work behaviour is justified. The adoption of a subjective ratings based performance appraisal instrument requires further justification, particularly in the inclusion of 'predictor-like' individual difference characteristics. Despite criticisms of this approach including the possibility that this method requires raters to draw upon theoretical constructs that they may have no

experience or knowledge of, (see for example, Armstrong, 2000), the advantage of this approach lies in the generalisability of the instrument across occupational and organisational boundaries. Potentially such an instrument could be used to assess infinite combinations of characteristics and abilities believed (and demonstrated through validation research) to be related to job performance.

10.4 The managerial personality

Development of biodata items to predict personality traits and motivational aspects of behaviour is not new (e.g. Kilcullen et al., 1995; Mael & Ashforth, 1995; Mael & Hirsch, 1993; Mount et al., 2000; Mumford et al., 1996; Wilkinson, 1995). The evidence suggest that construct-oriented scales demonstrate construct validity, generalisability across organisational and occupational groups, less decay over time than traditional empirical biodata analogues, and can predict aspects of performance including organisational loyalty and attrition rates. The current concurrent validity study represents a novel undertaking whereby analogues of critical thinking ability and extroversion and neuroticism were used to predict (concurrently) managerial performance using supervisor ratings and objective data as the criterion.

Critical thinking ability and the traits of extroversion and neuroticism appear to be inconsistently related or may have a non-linear relationship with managerial performance in the current sample. The sample size prevents meaningful analysis of the differential contribution of critical thinking, extroversion and neuroticism across different managerial functions. The selection of personality constructs in this study is rooted in the empirical findings outlined above and also in the theoretical value of Eysenck's (1957) model of

personality. The choice ultimately depends upon subscription to either a fundamentally descriptive approach (the five-factor model) or an explanatory view of personality (extroversion and stability). For those interested in the causes of behaviour then a two-factor model is more appropriate. For the present investigation, it is proposed that the latter has more value in terms of explaining why and how employees may become susceptible to stress, illness, and may offer greater scope for explanation of patterns of interaction in the work place.

The reported positive correlations between conscientiousness and managerial performance ratings from previous research (e.g. Gellatly, 1996) may be artifactual and it is certainly difficult to relate to underlying, explained and enduring 'permanent' traits. Similarly, it is difficult to untangle cause and effect relationships of within organisational citizenship behaviours of this nature and the appearance of traits such as 'agreeableness' and perhaps 'openness to experience'. However, the five-factor model of personality may not fit applicant populations as well as it has with volunteer populations (Schmitt and Ryan, 1993). The development of items that include five-factor items that relate directly to work content may provide more accurate measures of likely performance (indicating job knowledge as well as personality), future research must take account of this.

The supervisor providing the performance appraisal rating may consider the apparent conscientiousness of the employee as an aid to completing the appraisal. Conscientiousness and particularly perceived conscientiousness become 'predictor' rather than 'criterion' variables, while being utilised to produce evidence of predictive validity between two measures of the same construct.

Apparent conscientiousness may therefore operate via a 'halo' or 'horns' error mechanism and would function as a product of the limited information processing capacity of the rater (called upon to attend to various tasks across numerous situations, or not). A social psychological explanation may be employed in addition to this 'cognitive miser' approach. The research into the effects of liking and perceived similarity has uncovered the possibility that social impression management may enhance performance ratings. The image of conscientiousness, the apparent performance of important organisational citizenship behaviours, could all produce a positive correlation between conscientiousness and performance ratings (potentially supervisor, peer and self ratings could be influenced). This is likely to be true in situations where the performance appraisal proforma asks the rater to attend to the perceived level of conscientiousness demonstrated by the employee, as shown in the strong positive correlation between ratings of performance and perceived 'promotability' and apparent conscientiousness in this research.

The results of Hartrup, Connell & Wingate (1998) research into the prediction of sales performance provides supportive evidence for the possibility of an artifactual explanation for the positive correlations reported between job performance ratings and conscientiousness. Hartrup et al accept Borman & Motowidlo's (1993, 1997) distinction of task performance and contextual job performance and report that cognitive ability offers the best prediction of sales performance (controlling for conscientiousness) although it is not a good predictor of absenteeism or tardiness. Such behaviours may be summarised as the antithesis of organisational citizenship behaviour. This conclusion is supported by the present research in relation to the perceived conscientiousness of the managers rather than actual psychometric self-report measures. In fact, perceived conscientiousness is likely to offer a clearer picture of the social desirability of the trait of conscientiousness than self-

report measures. The question of the links between the appearance of this trait and a wider sociological construct of impression management provide interesting avenues for further, perhaps qualitative research. The possible contribution of Turner's (1987) self-categorisation theory, drawing upon the idea of a prototypical employee, may also provide some organisational answers to the micro-level questions of having a 'face that fits'.

Barker (2001) asserts that management differs from leadership, as the primary aim of management is the maintenance of structure and stability across the organisation. In contrast, leadership requires the 'dynamic exchange' between the collective will of the individuals within the organisation and the individual needs of those employees. In this way, leadership is oriented towards change and adaptation (Barker, 2001:491). In this model of management, perhaps the trait of conscientiousness would serve the manager well and the leader not so well. The distinction between management and leadership within occupational roles in practise is difficult however.

A reconsideration of Mischel's (1977, 1981) theory of personality with reference to 'person variables' can provide a means to understanding the incremental validity of biodata over traditional psychometric measures of personality traits. In summary, Mischel (1981) argues that behaviour is determined by a combination of the demands and restrictions of the situation and individual person variables. The low correlations reported between traditional personality measures and behaviour is said to be a result of the 'consistency paradox'. Mischel asserts that cross-situational consistency, or the extent to which an individual behaves the same way across different situations is unlikely. The appearance of cross-situational consistency emerges from the tendency to consider and describe prototypic behaviours when thinking about the personality or character of those

we know. In contrast, temporal consistency, consistency over time (but not situation), is more likely and would offer some support for a trait conceptualisation of personality as fairly stable or enduring traits over time. This model of personality offers a critical appraisal of the trait strategy of personality without necessarily dismissing the usefulness of traits for the description of traits or behavioural preferences over time. It is likely that a biographical life history inventory may capture stability over time more accurately than standard psychometric personality inventories as the validity and reliability of the biodata questionnaire approach typically relies on the previous experience and the robust finding that the best predictor of future behaviour is past behaviour (Wernimont & Campbell, 1968).

The possibility that ratings of performance are based upon prototypic judgements of personality (and indirectly of behaviour, and vice versa), so that with the possible exception of situations with few demand characteristics, job performance behaviour is unlikely to correlate above 0.3 (Mischel, 1977) with psychometric indicators of personality. Non-prototypic behaviour and performance may be overlooked so that the trait concept or in this instance, the job performance rating, may remain stable. This prototypicality argument could be integrated with an understanding of halo/horns error resulting in a 'cognitive bias' model of performance judgements such as those collected with regard to the perceived traits of conscientiousness and energy.

Organ (1994) asserts that organisational citizenships consist of a combination of personality traits and job attitudes. These are summarised as 'constructive, spontaneous, optional, non-compensated contributions...' (Organ, 1994:465). To date, there is no explicit evidence of the empirical validity of dispositional affectivity as a determinant or

predictor of organisational citizenship behaviours. The literature is peppered with the semantic leap between the five-factor (e.g. Bernardin, Cooke & Villanova, 2000; Mount & Barrick, 1998) 'trait' or factor of 'conscientiousness' and performance rating variance. At the same time, researchers concerned with the low correlations between personality measures and performance outcomes have tended to develop criterion measures capturing the same variance as the predictor to create larger (artifactual) correlations. Both ends of the validation continuum are predicting personality variance that may be associated with organisational citizenship behaviour.

Through the conversations with managerial job incumbents, motivation, enthusiasm, energy and conscientiousness appeared in several accounts of the ideal manager. In particular, one participant (P5) discusses the important of going 'beyond the call of duty' for managerial success (see section 9.). These characteristics appear to underline a sense of organisational citizenship behaviours of the type that Organ (1994) outlines.

Although Organ & Lingl (1995) report common variance across measures of agreeableness and conscientiousness facets of 'five-factor' model measures of personality, Organ (1994, 1997) concludes that organisational citizenship behaviours cannot be well predicted from personality variance alone. It is likely that prototypical organisational citizenship behaviours may look similar to implicit personality theories about the nature and appearance of conscientiousness.

Chapter 11 Conclusions

11.1 Limitations of the research

The managerial job incumbent sample involved in the main part of this investigation was largely male and middle-aged with a stable career history within the organisation (average age was 42 years and average length of time in the organisation was 10 years). The supervisors predicting the performance ratings were also white, middle-aged males for the most part. Ethnic minority groups were not represented in this part of the investigation at all. Despite this confounding of the criterion and concurrent validity study, several of the 'potential managers' involved in the pilot investigation represented EU and non-EU countries and reported having their pre-university education outside of the UK. This sample was the basis for the development of the biodata analogues, completing the biodata inventory, the EPI and the critical thinking appraisal test. The average length of time in the organisation suggests that the organisation is one characterised by fairly low levels of managerial turnover or attrition. It is possible that any conclusions drawn about critical thinking ability, extroversion and neuroticism for this sample may not generalise in smaller, private sector or 'faster-paced' organisations.

11.2 Limitations of scope and key assumptions

The fundamental assumption of this research lies in the belief that 'samples' of behaviour offer better prediction of future behaviour than 'signs' (Wernimont & Campbell, 1968). This assumption underlies the tradition of biographical data collection and is supported by

the impressive validity coefficients demonstrated through the development of empirically and rationally derived biodata.

A further assumption of this project is that the psychometric constructs under investigation exist and can be captured by the instruments used, namely, the Watson Glaser Critical Thinking Appraisal (Watson & Glaser, 1980) and the Eysenck Personality Inventory (Eysenck & Eysenck, 1963). Essentially, this project focuses upon the potential contribution that the psychometric constructs of critical thinking ability, extroversion and neuroticism may make in the prediction of managerial job performance. This research does not directly address the current controversy surrounding the issue of conscientiousness and managerial performance (e.g. Hough, 1992; Barrick, Mount & Strauss, 1993; Gellatly, 1996; and Hough, Ones & Viswesvaran, 1998).

Thirdly, the approach to performance ratings adopted in this research has necessitated a top-down, supervisor-driven approach and does not represent the reality of the Personal Planning Development process that is increasingly popular in human resource development practice (Armstrong, 2000). Sulsky and Balzer (1988) assert that there is no common definition of performance rating accuracy. They argue that traditionally, the accuracy of performance ratings has been calculated by comparing raters' performance ratings of a number of ratees on several performance dimensions with ratings from "expert raters", (expert ratings were obtained by computing pooled average scores). The problem of contamination of the criterion is endemic to this research and personnel psychology as a whole. The supervisor rating method was selected for this research following a literature review on supervisor, self and peer appraisal and the errors and bias associated with the traditional forms of these techniques. However, research utilising an interactive model of

performance management, perhaps using the repertory grid technique, could form a useful starting point for further discussion of underlying traits and abilities believed to promote or enhance managerial success. Such research may require a sample of managers that offer greater homogeneity than those in the present study and would certainly require a qualitative focus in methodology centred around interviews and focus groups such as that undertaken by Armstrong and Baron (1998) in their review of perceptions of performance management systems.

11.3 Suggestions for further research

A truly predictive validation study where actual managerial applicants complete the biodata inventories and psychometric measures would inform on the incremental validity of the analogue models and collect information relating to possible response distortion and motivated responding. Ultimately, this form of analysis could imbue the biodata inventory with greater utility when the construct-oriented approach to biodata development moves out of the developmental phase into pre-selection practice.

Further use of the current performance appraisal instrument across different organisations and managerial functions would provide information regarding the generalisability of the incremental validity and permit further cross-validation analyses.

The development of further biodata analogues predicting the success of varying occupational groups would provide additional optimistic evidence in support of this approach to biodata development. It is likely that development of analogues for more specifically defined occupational groups or categories, incorporating an empirical person

specification system (perhaps using current employee perceptions of the skills required for their work) as the basis for biodata, may produce more significant correlations between the individual differences constructs of interest and the performance outcomes collected but would inevitably reduce the generalisability of the findings.

Further work concerning motivation and job performance ('organisational citizenship behaviours') accounting for organisational culture, and incorporating qualitative analysis of current employee perceptions of performance and success with candidate perceptions of success would provide very interesting avenues for future biodata development. In particular, it is likely that the variable of 'career progress' consists of decisions about promotions and length of time at the organisation and is the result of probably more than one supervisor decision of worker effectiveness. To this end, the career progress findings may be considered an especially useful criterion (potentially less contaminated by intra-rater error) for further developmental research.

A key finding from this research must be a note of caution to all researchers working towards individual differences explanations of managerial job performance. An integrative approach combining personality, ability and social psychological explanations, such as the organisational citizenship behaviour model proposed by Organ (1997) among others, appears to offer the most productive way forward. Similarly, the need for collection of subjective and objective personnel data is underlined by the variance associated with both approaches in relation to interpersonal characteristics of the rater and ratee, and organisational economic factors.

It is likely that given Sternberg's (1997) compelling theoretical basis and collated evidence in support of the construct-validity of a triarchic model of managerial intelligence, such measures may produce better prediction of total performance rating and career progress. Sternberg asserts that managerial intelligence consists of analytical intelligence, practical and creative intelligence. Practical intelligence is said to consist of tacit knowledge conceptualised as 'action-oriented knowledge acquired without direct help from others, that allows individuals to achieve the goals they personally value' (Sternberg, 1997:483). Creative intelligence relates to the extent to which an individual is able to 'see problems in new ways and to escape the bounds of conventional thinking' (Sternberg, 1997:488). Biographical items designed to capture the dynamic model of 'managerial intelligence' proposed by Sternberg (1997) arte likely to offer incremental validity over measures of the triarchic abilities in the same way that biographical data offers incremental validity over traditional cognitive ability tests and is likely to provide greater criterion-related validity in the prediction of job performance ratings and career progress

11.4 Summary

This thesis outlines the research involved in the development of construct-oriented biodata for the prediction of managerial performance. In essence, the prediction that a construct-valid instrument will demonstrate incremental validity over the original measures of the constructs of critical thinking ability, extroversion, and neuroticism is supported. This demonstration supports the move across recent literature (Wilkinson, 1995; Mael & Hirsch, 1993; Mumford et al, 1996) towards construct-oriented or 'quasi-rational' biodata. It is beneficial to biodata scale development to base item generation upon the prediction of psychological constructs that are believed to underpin performance outcomes in order to

develop scales with acceptable internal consistency. It must be noted however, that given the low criterion-related validity of the psychological constructs under investigation in this research, the gains in incremental validity may appear to be more significant than they are. Tables 8.12, 8.13, and 8.14 illustrate the gain in validity from including the biodata analogue as a predictor against the performance outcomes, the original validity in all cases was small enough to make any % gain appear substantial.

The current results suggest that despite the low figures reported for current UK use of biodata, (Harvey-Cook, 2000), development of construct-oriented biodata for use in the pre-selection stage of selection can increase the utility of any existing selection programme though incremental validity gains and, through the measurement of psychological constructs chosen on the basis of a thorough person-specification procedure (Wilkinson van Zwanenberg, 1994) can increase understanding of the factors associated with managerial performance ratings on one hand, and objective indicators of success on the other.

Unit weighted and weighted item biodata analogue models demonstrate incremental validity over original psychometric test scores of critical thinking, extroversion and neuroticism for the prediction of total performance ratings and career progress. The primary hypothesis of the research is therefore supported.

Critical thinking, extroversion and neuroticism are not significantly correlated with total performance rating or career progress though the sample may suffer from restriction of range on the part of the personality scores (as a concurrent validity study, the possibility that extremes have been selected out prior to the experiment must be considered) and in

supervisor ratings of total performance as collected on a 6 point Likert scale. The components of supervisor ratings of managerial performance have been investigated and perceived conscientiousness and energy contribute much of the variance associated with overall performance ratings, suggesting the likelihood of halo error in the ratings and offering ground for a social psychological explanation of the results relating to this criterion.

Overall, the findings presented in this thesis suggest that further work in development of construct-oriented biodata analogues of stable individual differences may offer incremental validity and utility to pre-selection. In summary, construct-oriented biodata analogue models of critical thinking ability, extroversion and neuroticism, demonstrated incremental validity over the original psychometric measures of these constructs.

References

Allen, T.D., & Rush, M.C. (1998). The effects of organisational citizenship behaviour on performance judgements: A field study and a laboratory experiment. *Journal of Applied Psychology*, 83, 2, 247-260.

Allworth, E.G., & Hesketh, B. (1998, April). Generalisability of construct-oriented biodata scales in predicting adaptive performance. Paper presented at the 13th Annual Conference of the Society for Industrial and Organisational Psychology, Dallas, TX. Cited in Mount, M.K., Witt, L.A., Barrick, M.R. (2000). Incremental validity of empirically keyed biodata scales over GMA and the five factor personality constructs. *Personnel Psychology*, 53, 299-323.

American Psychological Association, American Educational Research Association, & National Council on Measurement in Education (Joint Committee). (1985) Standards for educational and psychological testing. Washington, D.C: American Psychological Association.

Anderson, N., & Shackleton, V. (1986). Recruitment and selection: A review of developments in the 80s. *Personnel Research*, 15, 4, 19-26.

Armstrong, M. (2000). *Performance management: Key strategies and practical guidelines*. Second edition. Kogan Page Ltd, London.

Armstrong, M., & Baron, A. (1998). *Performance management: The new realities*. Institute of Personnel and Development, London.

Arvey, R. D. (1979). Unfair discrimination in the employment interview: Legal and psychological aspects. *Psychological Bulletin*, 86, 736-765.

Arvey, R. D., & Campion, J.E. (1982). The employment interview: A summary and review of recent aspects. *Psychological Bulletin*, 92, 436-446.

Asher, J. (1972). The biographical item: Can it be improved? *Personnel Psychology*, 25, 251-269.

Ashforth, B.E., & Mael F. (1989). Social identity theory and the organization. *Academy of Management Review*, 14, 1, 20-39.

Astley, G. (1984). Subjectivity, sophistry and symbolism in management science. *Journal of Management Studies*, 21, 3, 259-71.

Baehr, M. E., & Orban, J. A. (1989). The role of intellectual abilities and personality characteristics In *Assessment and selection in organizations: Methods and practice for recruitment and appraisal.* (pp. 275-309). P. Herriot (Ed.). Chichester, England: John Wiley and Sons Ltd.

Baehr, M.E., & Williams, G.B. (1967). Underlying dimensions of personal background data and their relationship to occupational classification. *Journal of Applied Psychology*, 51, 6, 481-490.

Barclay, J.M. (2001). Improving selection interviews with structure: organizations use of 'behavioural' interviews. *Personnel Review*, 30, 1, 81-101.

Barge, H.M., & Hough, L. M. (1986). Utility of biographical data for predicting job performance. In Hough, L. M. (Ed), Literature review: Utility of temperament, biodata, and interest assessment for predicting job performance. (Research note 88-02). Alexandria, VA: US Army Research Institute. Cited in Mael, F.A., & Ashforth, B. E. (1995). Loyal from day one: Biodata organisational identification and turnover among newcomers. *Personnel Psychology*, 46, 719-738.

Barker, R.A. (2001). The nature of leadership. *Human Relations*, 54, 4, 469-494.

Barrett, P., & Hutton, R. (2000). Psychometrics and personality questionnaires. *Selection and Development Review*, 16, 2, 5-9.

Barrick, M.R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1-26.

Baruch, Y. (1998). Self-performance appraisal vs direct-manager appraisal: A case of congruence. *Journal of Managerial Psychology*, 11, 6, 1-12.

Bass, B.M. (1985). *Leadership and performance beyond expectations*. New York, Free Press.

Beason, G., & Belt, J. A. (1976). Verifying applicants' backgrounds. *Personnel Journal*, 55, 345-348.

Bem, S.L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, 42, 155-162.

Bendig, A.W. (1963). The relation of temperament traits of social extroversion and emotionality to vocational interests. *Journal of General Psychology*, 69, 311-318.

Bennett, G.K., Seashore, H.G., & Wesman, A. (1974). *Differential aptitude tests*. New York: Psychological Corporation.

Bennis, W., & Nanus, B. (1985). *Leaders*. New York: Harper & Row.

Ben-Shakhar, B., Bar-Hillel, M., Bilu, Y., Ben-Abba, E., & Flug, A. (1980). Can graphology predict occupational success? Two empirical studies and some methodological ruminations. *Journal of Applied Psychology*, 71,645-53.

Bergwerk, J. (1991). Recruitment and selection for company culture. *Personnel Review,* 20, 1, 9-16.

Bernardin, H.J. (1978). Effects of rater training on leniency and halo errors in student ratings of instructors. *Journal of Applied Psychology*, 63, 3, 301-308.

Bernardin, H.J., & Walter, C.S. (1977). Effects of rater training and diary-keeping on psychometric error in ratings. *Journal of Applied Psychology*, 62, 1, 64-69.

Bernardin, H.J., Cooke, D.K., & Villanova, P. (2000). Conscientiousness and agreeableness as predictors of rating leniency. *Journal of Applied Psychology*, 85, 2, 232-234.

Berry, L. M., & Houston, J.P. (1993). *Psychology at work: An introduction to industrial and organisational psychology.* USA: Brown & Benchmark.

Bethell-Fox, C. E. (1989). Psychological testing. *In Assessment and selection in organisations: Methods and practice for recruitment and appraisal.* P. Herriot (Ed.). Chichester, England: John Wiley and Sons Ltd.

Bleisener, T. (1996). Methodological moderators in validating biographical data in personnel selection. *Journal of Occupational and Organisational Psychology*, 69, 107-120.

Blinkhorn, S., & Johnson, C. (1990). The insignificance of personality testing. *Nature*, 348, 671-62.

Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin,* 117, 2, 187-215.

Blood J.W. (Ed.) *The personnel job in a changing world.* New York: American Management Association.

Bommer, W.H., Johnson, J.L., Rich, G. A., Podsakoff, P.M., & Mackenzie, S. B. (1995). On the interchangeability of objective and subjective measures of employee performance: A meta-analysis. *Personnel Psychology*, 48, 587-605.

Borman, W.C., & Brush, D. H. (1993). More progress toward a taxonomy of managerial performance requirements. *Human Performance*, 6, 1, 1-21.

Borman, W.C., Hanson, M. A., Oppler, S.H., Pulakos, E.D., & White, L. A. (1993). Role of early supervisory experience in supervisor performance. *Journal of Applied Psychology*, 78, 443-449.

Boyatzis, R. E. (1982). *The competent manager: A model for effective performance.* New York: John Wiley & Sons.

Brammer, I. (1992). Selection: Rear-view mirror or crystal ball? *Recruitment, Selection and Retention*, 1, 1, 20-24.

Bray, D., Campbell, R., & Grant, D. (1974). *Formative years in business: A long term AT&T study of managerial lives*. New York: Wiley.

Bridges, A. (1984). Assessment centres: their use in industry in Great Britain. *Psychological Monographs*, 80, (17, Whole No. 625).

Briggs, S. R. (1989). The optimal level of measurement for personality constructs. In D. M. Buss and M.C. Cantor (Eds.). *Personality Psychology: Recent trends and emerging directions*. New York: Springer-Verlag.

Briggs, S. R. (1992). Assessing the five-factor model of personality description. *Journal of Personality*, 60, 253-293.

Brown, S.H. (1994). Validating biodata. In G.S. Stokes, M.D. Mumford, & W.A. Owens (Eds.) *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction*. (pp. 199-236). Palo Alto, CA, USA: Cpp Books.

Brush, D.H., & Owens, W.A. (1979). Implementation and evaluation of an assessment classification model for manpower allocation. *Personnel Psychology*, 3, 369-383.

Cames, I., Vinnicombe, S., & Singh, V. (2001). Personal conceptions of the successful manager. *Women in Management Review*, 16, 3, 108-118.

Campbell, C.H., Ford, P., Rumsey, M. G., Pulakos, E. D., Borman, W.C., Felker, D. B., De Vera, M. V., & Riegelhaupt, B.J. (1990). Development of multiple job performance measures in a representative sample of jobs. *Personnel Psychology*, 43, 277-300.

Campbell, J. P. (1990). Modelling the performance prediction problem in industrial and organisational psychology. In M. Dunnette, and L.M. Hough, (eds.), *Handbook of industrial and organisational psychology*. Cambridge, Mass., Blackwell.

Campbell, J.P. (1991). Alternative models of job performance and their implications for selection and classification. In M.D. Dunnette & L. M. Hough (Eds). *The handbook of industrial and organisational psychology* (pp. 33 - 51). Palo-Alto: Consulting Psychologists Press.

Campbell, J.P., Dunnette, M.D., Arvey, R. D., & Hellervik, L.V. (1973). The development and evaluation of behaviourally based rating scales. *Journal of Applied Psychology*, 57, 1, 15-22.

Campbell, J.P., Dunnette, M.D., Lawler, E. E., & Weick, K. E. Jr. (1970). *Managerial behaviour, performance, and effectiveness*. New York: McGraw-Hill.

Cardy, R.L., & Kehoe, J.F. (1984). Rater selective attention ability and appraisal effectiveness: The effect of a cognitive style on the accuracy of differentiation among ratees. *Journal of Applied Psychology*, 69, 4, 589-594.

Cardy, R.L., Dobbins, G.H., & Carson, K.P. (1995). TQM and HRM: Improving performance appraisal research, theory, and practice. *Canadian Journal of Administrative Sciences*, 12, 2, 106-115.

Carlson, S. (1951). *Executive behaviour: A study of workload and the working methods of managing directors*. Stockholm: Stromberg.

Carlson, K.D., Scullen, S.E., Schmidt, F.L., Rothstein, H., & Erwin, F. (1999). Generalisable biographical data validity can be achieved without multi-organisational development and keying. *Personnel Psychology*, 52, 731-755.

Carrier, M.R., Dalessio, A. T., & Brown, S.H. (1990). Correspondence between estimates of content and criterion-related validity values. *Personnel Psychology*, 43, 85-100.

Carroll, S. J. Jr., & Taylor, W.H. (1968). A study of the validity of a self-observational central signalling method of work sampling. *Personnel Psychology*, 21, 359-364.

Carroll, S.J., & Nash, A. N. (1972). Effectiveness of a forced choice reference check. *Personnel Administration*, 35, 42-146.

Carroll, S.J., Jr., & Gillen, D. J. (1987). Are the classical functions useful in describing managerial work? *Academy of Management Review*, 12, 1, 38-51.

Cascio, W.F. (1978). *Applied Psychology in Personnel Management*. Reston, CA: Reston Publishing Company.

Cascio, W.F. (1991). *Applied psychology in personnel management*. (4th Ed.). Englewood Cliffs, NJ: Prentice-Hall.

Cascio, W.F. (1995). Whither industrial and organisational psychology in a changing world of work? *American Psychologist*, 50, 11, 928-939.

Cascio, W.F., & Phillips, N. F. (1979). Performance testing: A rose among thorns? *Personnel Psychology*, 32, 751-766.

Cascio, W.F., & Sibley, Utility of the assessment centre as a selection device. *Journal of Applied Psychology*, 64, 107-118.

Cascio, W.F., & Valenzi, E. R. (1978). Relations among the criteria of police performance. *Journal of Applied Psychology*, 63, 22-28.

Cawley, B. D., Keeping, L. M., & Levy, P.E. (1998). Participation in the performance appraisal process and various employee reactions: A meta-analytic review of field investigations. *Journal of Applied Psychology*, 83, 4, 615-633.

Chen, X-P, Hui, C., & Sego, D.J. (1998). The role of organisational citizenship behaviour in turnover: Conceptualisation and preliminary tests of key hypotheses. *Journal of Applied Psychology*, 83, 6, 922-931.

Childs, A. & Klimoski, R.J. (1986). Successfully predicting career success: An application of the biographical inventory. *Journal of Applied Psychology*, 71, 1, 3-8.

Chow, W. S., & Luk, V.W.M. (1996). Management in the 1990s: A comparative study of women managers in China and Hong Kong. *Journal of Managerial Psychology*, 11, 1, 24-36.

Clark, T. (1992). Management selection by executive recruitment consultancies: A survey and explanation of selection methods. *Journal of Management Psychology*, 7, 6, 3-10.

Cleveland, J.N., & Landy, F.J. (1981). The influence of rater age and ratee age on two performance judgements. *Personnel Psychology*, 34, 19-29.

Cohen, B.M., Moses, J.L., & Byham, W.C. (1974). *The validity of assessment centres: A literature review*. Pittsburgh: Development Dimensions Press.

Collins, J. (1998). Conscientiousness: Is that all there is? Society for Industrial and Organisational Psychology, *The Industrial Psychologist* (TIP), 36, 2. Division of the American Psychological Association Organisational Affiliate of the American Psychological Society.

Connerley, M.L, Mael, F.A., & Morath, R.A. (1999). 'Don't ask-Please tell': Selection privacy from two perspectives. *Journal of Occupational and Organisational Psychology*, 72, 405-422.

Conway, J.M., Jako, R.A., & Goodman, D.F. (1995). A meta-analysis of inter-rater and internal consistency reliability of selection interviews. *Journal of Applied Psychology*, 80, 5, 565-579.

Cook, M. (1993). *Personnel selection and productivity.* Second Edition. Chichester, England: John Wiley and Sons.

Cook, M. (1998). *Personnel selection: Adding value through people.* Third edition. Chichester, England: John Wiley and Sons.

Cook, T., & Emler, N. (1999). Bottom-up versus top-down evaluations of candidates' managerial potential: an experimental study. *Journal of Occupational and Organisational Psychology*, 72, 423-439.

Cooper, D., & Robertson, I. T. (1995). *The Psychology of Personnel Selection: A Quality Approach*. Routledge, London.

Cooper, W.H. (1981). Ubiquitous halo. *Psychological Bulletin*, 90, 218-244.

Cortina, J.M., Doherty, M. L., Schmitt, N., Kaufman, G., & Smith, R.G. (1992). The "Big Five" personality factors in the IPI and the MMPI: Predictors of police performance. *Personnel Psychology*, 45, 119-140.

Crawley, B., Pinder, R., & Herriot, P. (1990). Assessment Centre Dimensions, Personality and Aptitudes. *Journal of Occupational Psychology*, 63, 211-216.

Cronbach, L.J. (1990). *Essentials of psychological testing*. Fifth Edition. Harper & Row Publishers, Inc., New York.

Cronbach, L.J., & Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.

Crowne, D. P, & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology,* 24, 349-354.

Dale, E. (1965). *Management: Theory and practice.* New York: Wiley.

Dalton, M. (1959). *Men who manage.* New York: Wiley.

Day, D.V., & Silverman, S. B. (1989). Personality and job performance: Evidence of incremental validity. *Personnel Psychology*, 42, 25-36.

Dean, M.A., & Russell, C.J. (1998, April). A comparison of g and biodata criterion-related validity in a sample of air traffic controllers. Paper presented at the 13th Annual Conference of the Society for Industrial and Organisational Psychology, Dallas, TX. Cited in Mount, M.K., Witt, L.A., Barrick, M. R. (2000). Incremental validity of empirically keyed biodata scales over GMA and the five factor personality constructs. *Personnel Psychology,* 53, 299-323.

Deluga, R.J. (1998). Leader-member exchange quality and effectiveness ratings: The role of subordinate-supervisor conscientiousness similarity. *Group and Organization Management*, 23, 2, 189-216.

DeNisi, A. S., Cafferty, T.P., & Meglino, B.M. (1984). A cognitive view of the performance appraisal process: A model and research propositions. *Organizational Behaviour and Human Performance*, 33, 360-396.

Digman, J.M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology,* 41, 417-440.

Dobbins, G.H., Cardy, R.L., & Truxillo, D.M. (1988). Effect of ratee sex and purpose of appraisal on the accuracy of performance evaluations. *Basic and Applied Social Psychology*, 7, 225-241.

Dobson, P. (1989). Reference Reports. In. P. Herriot (Ed.). *Assessment and Selection in Organizations: Methods for recruitment and appraisa*l (pp455-488). Chichester, England: John Wiley and Sons.

Dobson, P. (1989). Self and Peer Assessment. In P. Herriot (Ed.) *Selection and Assessment in Organizations: Methods and Practice for Recruitment and Appraisal* (pp. 421-433). Chichester, England: John Wiley and Sons.

Drakeley, R. (1989). Biographical data. In *Handbook of Assessment and selection in organizations: Methods and practice for recruitment and appraisal. P. Herriot*

(ed). John Wiley and Sons, Ltd.

Drakeley, R. J., Herriot, P., & Jones, A. (1988). Biographical data, training, success, and turnover. *Journal of Occupational Psychology*, 61, 145-152.

Dreher, G.F., & Bretz, R. D., Jr. (1991). Cognitive ability and career attainment: Moderating effects of early career success. *Journal of Applied Psychology*, 76,392-397.

Drucker, P.F. (1977). *Management.* London: Pan.

Dunn, W. S., Mount, M. K., Barrick, M.R., & Ones, D.S. (1995). Relative importance of personality and general mental ability in managers' judgements of applicant qualifications. *Journal of Applied Psychology,* 80, 4, 500-509.

Dunnette, M.D. (1962). A note on the criterion. *Journal of Applied Psychology*, 47, 251-254.

Dunnette, M.D. (1972). *Validity study results for jobs relevant to the petroleum refining industry.* American Petroleum Institute.

Dunnette, M.D., McCartney, J., Carlson, H. C., & Kirchner, W. K. (1962). A study of faking behaviour on a forced-choice self-description checklist. *Personnel Psychology*, 15, 13-24.

Eberhardt, B.J., & Muchinsky, P.M. (1982). An empirical investigation of the factor stability of Owens' Biographical Questionnaire. *Journal of Applied Psychology*, 67, 2, 138-145.

Edwards, A.G., & Armitage, P. (1992). An Experiment to test the discriminatory ability of Graphologists. *Personality and Individual Differences*, 13 (1) 69-74.

Elkins, T.J., & Phillips, J.S. (2000). Job context, selection decision outcome, and the perceived fairness of selection tests: Biodata as an illustrative case. *Journal of Applied Psychology*, 85, 3, 479-484.

Elkins, T.J., & Phillips, J.S. (2000). Job context, selection decision outcome, and the perceived fairness of selection tests: Biodata as an illustrative case. *Journal of Applied Psychology*, 85, 3, 479-484.

Entwhistle, N. J., & Cunningham, S. (1971). Neuroticism and school attainment - a linear relationship? In H.J. Eysenck (Ed.) *Readings in extroversion and Introversion: 2 fields of application.* London: C. Tingling & Co.

Eysenck, H.J. (1967). Personality patterns in various groups of businessmen. *Occupational Psychology*, 41, 249-50.

Eysenck, H.J. (1989). Review of the California Personality Inventory. In J.C. Conoley and J.J. Kramer (Eds.), *The Tenth Mental Measurements Yearbook.* Lincoln: Buros Institute of Mental Measurements of the University of Nebraska.

Eysenck, H.J. (1991). Dimensions of personality: 16: 5 or 3? Criteria for a taxonomic paradigm. Personality and Individual Differences, 12, 8, 773-790.

Eysenck, H.J.& Eysenck, M.W. (1985). *Personality and individual differences: A natural science approach.* Plenum Press: New York.

Eysenck, H.J., & Eysenck, S.B.G. (1963). *Eysenck Personality Inventory.* London: University of London Press

Eysenck, H.J., & Eysenck, S.B.G. (1964). *Manual of the Eysenck Personality Inventory.* Hodder and Stoughton, England.

Ezell, H.F., Odewahn, C.A., & Sherman, J.D. (1982). Women entering management: Differences in perceptions of factors influencing integration. *Group and Organization Studies*, 7, 243-253.

Fayol, H. (1949). *General and industrial management.* (C. Storrs, Trans.). London, England: Pitman.

Feldman, J.M. (1981). Beyond attribution theory: Cognitive processes in performance appraisal. *Journal of Applied Psychology*, 66, 2, 127-148.

Feltham, R. T. (1988). Assessment centre decision making: Judgemental vs. mechanical. *Journal of Occupational Psychology*, 61, 237-241.

Feltham, R.T. (1988). Validity of a police assessment centre: A 1-19-year follow-up. Journal of Occupational Psychology, 61, 129-144.

Feltham, R.T. (1989). Assessment centres. In P. Herriot (Ed.). *Assessment and selection in organizations: Methods for recruitment and appraisal* (pp. 401-420). Chichester, England: John Wiley and Sons.

Fiedler, F. E. (1996). Research on leadership selection and training: One view of the future. *Administrative Science Quarterly*, 41, 2, 241-251.

Findlay, H.M., Giles, W.F., & Mossholder, K.W. (2000). Performance appraisal process and system facets: relationships with contextual performance. *Journal of Applied Psychology*, 85, 4, 634-640.

Fine, S.A., & Getake, M. (1995). *Benchmark tasks for job analysis: A guide for functional job analysis (FJA) scales.* New Jersey: Lawrence Erlbaum Associates.

Fine, S.A., & Wiley, W.W. (1974). An introduction to job analysis. In, E.A. Fleishman (Ed.) *Studies in personnel and industrial psychology* (3rd ed.) (pp. 6-13). New York: Dorset Press.

Finkle, R.B. (1976). Managerial Assessment Centres. In M. D. Dunnette (Ed.) *Handbook of Industrial and Organizational Psychology* (pp. 861-888). Rand McNally College Publishing Company, Chicago.

Flanagan, J.C. (1946). The experimental validation of a selection procedure. *Educational and Psychological Measurement*, 6, 445-466.

Flanagan, J.C. (1954). The critical incident technique. *Psychological Bulletin*, 51, 327-358.

Fletcher, C. (1991). Candidates' reactions to assessment centres and their outcomes: A longitudinal study. *Journal of Occupational Psychology*, 64, 117-127.

Fox, A. (1974). *Beyond contract.* London: Faber.

Furnham, A. (1997). *The psychology of behaviour at work: The individual in the organization.* Psychology press Publishers, Sussex, UK.

Furnham, A., & Stringfield, P. (1998). Congruence in job-performance ratings: A study of 360 degrees feedback examining self, manager, peers, and consultant ratings. *Human Relations*, 51, 4, 517-531.

Gattiker, U.E., & Larwood, L. (1989). Career success, mobility and extrinsic satisfaction of corporate managers. *Social Science Journal*, 26, 75-92.

Gaugler, B. B., & Thornton, G.C.,III. (1989). Number of assessment centre dimensions as a determinant of assessor accuracy. *Journal of Applied Psychology*, 74, 611-618.

Gaugler, B. B., Rosenthal, D. B., Thornton III, G.C., & Bentson, C. (1987). Meta-analysis of assessment centre validity. *Journal of Applied Psychology Monograph*, 72, 3, 493-511.

Gellatly, I. R. (1996). Conscientiousness and task performance: Test of a cognitive process model. *Journal of Applied Psychology*, 81, 5, 474-482.

Gellatly, I.R., Paunonen, S.V., Meyer, J.P., Jackson, D.N., & Goffin, R. D. (1991). Personality, vocational interest, and cognitive predictors of managerial job performance and satisfaction. *Personality and Individual Differences*, 12, 15-31.

Gersein, M., & Reisman, H. (1983). Strategic selection: Matching executives to business conditions. *Sloan Management Review*, Winter, 33-47.

Ghiselli, E. E. (1956). Dimensional problems of criteria. *Journal of Applied Psychology*, 40, 1 - 4.

Ghiselli, E. E. (1966). The validity of occupational aptitude tests. New York: Wiley.

Ghiselli, E. E. (1971). *Explorations in managerial talent.* Pacific Palisades, CA: Goodyear.

Ghiselli, E. E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psychology*, 26, 461-477.

Ghiselli, E. E., & Brown, C.W. (1955). *Personnel and industrial psychology* (2nd Ed.). New York: McGraw Hill.

Gill, D. (1980). *Selecting managers: How British industry recruits.* IPM, BIM, London.

Goffin, R.D., Rothstein, M.G., & Johnston, N.G. (1996). Personality testing and the assessment centre: Incremental validity for managerial selection. *Journal of Applied Psychology*, 81, 6, 746-756.

Goldberg, L.R. (1994). Basic research on personality structure: Implications of the emerging consensus for applications to selection and classification. In: M. G. Rumsey, C. B. Walker, & J.H. Harris (Eds.) *Personnel selection and classification* (pp. 247-259). New Jersey: Lawrence Erlbaum Associates.

Goldsmith, D.B. (1922). The use of personal history blanks as a salesmanship test. *Journal of Applied Psychology*, 6, 149-155.

Goldstein, H.W., Yusko, K.P., Braverman, E.P., Smith, D.B., & Chung, B. (1998). The role of cognitive ability in the subgroup differences and incremental validity of assessment centre exercises. *Personnel-Psychology,* 51, 2, 357-374.

Gratton, L. (1989). The work of the manager. In P. Herriot (Ed.) *Assessment and Selection in Organisations: Methods for recruitment and appraisal.* Chichester, England: John Wiley and Sons.

Greenberg, J. (1986). Determinants of perceived fairness of performance evaluations. *Journal of Applied Psychology*, 71, 2, 340-342.

Grimsley, G., & Jarrett, H.F. (1973). The relation of past managerial achievement to test measures obtained in the employment situation: Methodology and results. *Personnel Psychology*, 26, 31-48.

Grimsley, G., & Jarrett, H.F. (1975). The relation of past managerial achievement to test measures obtained in the employment situation: Methodology and results II. *Personnel Psychology*, 28, 215-231.

Guilford, J.P. (1959). *Personality.* New York: McGraw Hill.

Guion, R.M, & Gottier, R.F. (1966). Validity of personality measures in personnel selection. *Personnel Psychology,* 18, 135-164.

Guion, R.M. (1965). *Personnel testing.* New York: McGraw-Hill.

Guion, R.M. (1976). Recruitment, selection, and job placement. In M.D. Dunnette (Ed.) *Handbook of industrial and organizational psychology.* (pp. 777-828). Chicago; Rand McNally College Publishing Company.

Guion, R.M. (1983). Comments on Hunter. In F. Landy, S. Zedeck, & J.Cleveland (Eds.), *Performance measurement and theory* (pp267-275). Hillsdale, NJ: Erlbaum.

Guion, R.M., & Gottier, R.F. (1965). Validity of personality measures in personnel selection. *Personnel Psychology*, 18, 135-164.

Gulick, L.H. (1937) Notes on the theory of organization. In L.H. Gulick & L.F. Urwick (Eds.), Paper on the science of administration (pp3-20). New York: Columbia University Press. Cited in W.C. Borman & D.H. Brush (1993). More progress toward a taxonomy of managerial performance requirements. *Human performance*, 6, 1, 1-21.

Gulliford, R.A. (1991). The role of personality in assessing managerial potential. *Personnel Review*, 20, 1, 25-31.

Gunter, B., Furnham, A., & Drakeley, R. (1993). *Biodata: Biographical indicators of business performance.* Routledge; London.

Hales, C.P. (1986). What do managers do? A critical review of the evidence. *Journal of Management Studies*, 23, 1, 88-124.

Handyside, J. D., & Duncan, D.C. (1954). Four years later: A follow-up of an experiment in selecting supervisors. *Occupational Psychology*, 28, 9-23.

Harrell, T.W. (1972). High earning MBAs. *Personnel Psychology*, 25, 322-326.

Harris, M. M., Smith, D. E., Champagne, D. (1995). A field study of performance appraisal purpose: Research versus administrative based ratings. *Personnel Psychology*, 48, 151-160.

Hartigan, J. A., & Wigdor, A. K. (1989). *Fairness in employment testing*. Washington, DC: National Academy Press.

Hartle, F (1995). *Transforming the management process*. Kogan Page Ltd, London.

Harvey, R. J. (1991). Job analysis. In M.D. Dunnette & L. M. Hough (Eds). *The handbook of industrial and organisational psychology* (pp. 71-163). Palo-Alto: Consulting Psychologists Press.

Harvey-Cook, J. (2000). Demystifying biodata: A description of a practical application. *Selection and Development Review*, 16, 3, 15-19.

Hattrup, K., O'Connell, M. S., Wingate, P. H. (1998). Prediction of multidimensional criteria: Distinguishing task and contextual performance. *Human Performance*, 11, 4, 305-319.

Hayes, J., Rose-Quirie, A., & Allinson, C.W. (2000). Senior managers' perceptions of the competencies they require for effective performance: implications for training and development. *Personnel Review*, 29, 1, 92-105.

Hedge, J.W., & Kavanagh, M.J. (1988). Improving the accuracy of performance evaluations: Comparison of three methods of performance appraiser training. *Journal of Applied Psychology*, 73, 68-73.

Heim, A.W. (1967). *AH4 Group test of general intelligence*. Slough: National Foundation for Educational Research.

Heim, A.W. (1968). *Manual for the A5 Group Test of High Grade intelligence*. Slough: National Foundation for Educational Research.

Hemphill, J.K. (1959). Job descriptions for executives. *Harvard Business Review,* 37, 55-67.

Hemphill, J.K. (1960) Dimensions of executive positions. (Research monograph No. 9). Ohio State University, Bureau of Business Research. (Cited in Rose, S., & Baydoun, R. (1995). Content validity: A neglected strategy for developing managerial selection tests? *Current Psychology: Research & Reviews*, 14, 2, 138-151.).

Herold, D. M., Liden, R.C., & Leatherwood, M. L. (1987). Using multiple attributes to assess sources of performance feedback. *Academy of Management Journal,* 30, 826-835.

Herriot, P. (1989). The selection interview. In P. Herriot (Ed.). *Assessment and selection in organisations: Methods for recruitment and appraisal* (pp433-238). Chichester: John Wiley and Sons.

Herzberg, F. (1954). Temperament measures in industrial selection. *Journal of Applied Psychology,* 38, 81-84.

Hesketh, B., & Robertson, I. T. (1993). Validating personnel selection: A process model for research and practice. *International Journal of Selection and Assessment*, 1, 1, 3-17.

Hinrichs, J. R. (1978). An eight year follow-up of a management assessment centre. *Journal of Applied Psychology*, 63, 5, 596-601.

Hirsch, W., & Bevan, S. (1991). Managerial Competences and Skill Languages. In M. Silver (Ed.). *Competent to Manage*. England; Routledge.

Hobson, C. J., Mendel, R.M., & Gibson, F. W. (1981). Clarifying performance appraisal criteria. Organizational Behaviour and Human Performance, 28, 164-188.

Hochwarter, W.A., Witt, L.A., & Kacmar, K.M. (2000). Perceptions of organisational politics as a moderator of the relationship between conscientiousness and job performance. *Journal of Applied Psychology*, 85, 3, 472-478.

Hochwarter, W.A., Witt, L.A., & Kacmar, K.M. (2000). Perceptions of organisational politics as a moderator of the relationship between conscientiousness and job performance. *Journal of Applied Psychology*, 85, 3, 472-478.

Hogan, J. B. (1994). Empirical keying of background data measures. In G.S. Stokes, M.D. Mumford, and W.A. Owens (Eds.). *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction.* (pp. 69-107). Palo Alto, CA, USA: Cpp Books.

Hogan, R., Carpenter, B.N., Briggs, S. R., & Hansson, R.O. (1985). Personality assessment and personnel selection. In J. Bernardin and D. Bownass (Eds.) *Personality assessment in organisations*. New York: Praeger.

Hollander, E. P. (1965). Validity of peer nominations in predicting a distance performance criterion. *Journal of Applied Psychology*, 49, 434-438.

Hollon, C. J. (1983). Machiavellianism and managerial work attitudes and perceptions. *Psychological Reports*, 52, 432-434.

Holzbach, R.L. (1978). Rater bias in performance ratings: Superior, self-, and peer ratings. *Journal of Applied Psychology*, 63, 579-588.

Horvath, M., Ryan, A. M., & Steirwalt, S.L. (2000). The influence of explanations for selection test use, outcome favourability, and self-efficacy on test-taker perceptions. *Organisational Behaviour and Human Decision Processes*, 83, 2, 310-330.

Hough, C. M. (1992). The Big Five personality variables-construct: Confusion-confusion: Description versus prediction. *Human Performance*, 5, 139-155.

Hough, L. & Paullin, C. (1994). Construct-oriented scale construction: The rational approach. . In G.S. Stokes, M.D. Mumford, and W.A. Owens (Eds.). *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction*. (pp. 109-145). Palo Alto, CA, USA: Cpp Books.

Hough, L. M., Eaton, N.K., Dunnette, M.D., Kamp, J. D., & McCloy, R.A. (1990). Criterion-related validates of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology,* 75, 5, 581-595.

Hovland, C.I., & Wonderlic, E.F. (1939). Prediction of success by a standardised interview. *Journal of Applied Psychology*, 23, 537-546.

Howard, A., & Bray, D. (1988). *Managerial lives in transition: Advancing age and changing times.* New York: Guilford Press.

Huck, J. R., & Bray, D.W. (1976). Management assessment centre evaluations subsequent job performance of white and black females. *Personnel Psychology*, 29, 13-30.

Huck, J. R., & Russell, C. J. (1989). *Assessment centre validity in a continuing education setting.* Presented at the International Institute of Research, London, England.

Huck, J.R. (1973). Assessment centres: a review of the external and internal validities. *Personnel Psychology*, 26, 2, 191-212.

Huffcutt, A. I., & Arthur, W., Jr. (1994). Hunter and Hunter revisited: Interview validity for entry-level jobs. *Journal of Applied Psychology*, 79, 184-190.

Hull, C. L. (1928). *Aptitude testing.* Harrap, London.

Hunter, J.E. (1983). A causal analysis of cognitive ability, job knowledge, job performance, and supervisory ratings. In F. Landy, S. Zedeck, & J. Cleveland (Eds.), *Performance measurement and theory* (pp257-266). Hillsdale, NJ: Erlbaum.

Hunter, J.E., & Hunter, R.F. (1981). *The validity and utility of alternate predictors of job performance.* Unpublished manuscript.

Hunter, J.E., & Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96, 1, 72-98.

Hunter, J.E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-analysis: Cumulating research findings across studies.* Beverly Hills, California: Sage.

Ilgen, D. R., & Barnes-Farrell, J.L. (1984). *Performance planning and evaluation.* Chicago: SRA.

Ilgen, D. R., & Feldman, J.M. (1983). Performance appraisal: A focus process. *Research in Organizational Behaviour*, 5, 141-197.

Jackson, D., & Rothstein, M. (1993). Evaluating personality testing in personnel selection. *The Psychologist: Bulletin of The British Psychological Society*, 6, 8-11.

Jackson, D.N. (1990). Quantitative perspectives on the personality-job performance relationship. *The Score*, 13, 3-6.

James, L.R. (1973). Criterion models and construct validity for criteria. *Journal of Applied Psychology*, 80, 75-83.

Janz, T. (1982). Initial comparisons of patterned behaviour description interviews versus unstructured interviews. *Journal of Applied Psychology,* 67, 577-580.

Jenson, A. R. (1969). *Genetics and education.* London: Methuen.

Johnson, C., & Blinkhorn, (1994). Job performance and personality test validities: Desperate measures. *The Psychologist: Bulletin of The British Psychological Society*, 8, 167-170.

Jones, A., & Harrison, E. (1982). Prediction of performance in initial officer training using reference reports. *Journal of Occupational Psychology*, 55, 35-42.

Judge, T. A., & Bretz, R. D. (1994). Political influence behaviour and career success. *Journal of Management,* 20, 43-65.

Judge, T. A., Cable, D. M., Boudreau, J.W., & Bretz, R. D. (1995). An empirical investigation of the predictors of executive career success. *Personnel Psychology*, 48, 485-519.

Judge, T.A., & Bono, J.E. (2000). Five –factor model of personality and transformational leadership. *Journal of Applied Psychology*, 85, 5, 751-765.

Kane, J. S., & Lawler, E. E., III. (1978). Methods of peer assessment. *Psychological Bulletin*, 85, 555-586.

Karas, M & West, J.A. (1998, April). An explicit, construct-oriented approach to the development of rational-empirical biodata. Paper presented at the 13th Annual Conference of the Society for Industrial and Organisational Psychology, Dallas, TX. Cited in Mount, M.K., Witt, L.A., Barrick, M.R. (2000). Incremental validity of empirically keyed biodata scales over GMA and the five factor personality constructs. *Personnel Psychology*, 53, 299-323.

Katz, R.L. (1974). Skills of an effective administrator. *Harvard Business Review*, 52, 1, 90-102.

Kenny, D. A., & Zaccaro, S.J. (1983). An estimate of variance due to traits in leadership. *Journal of Applied Psychology*, 68, 678-685.

Kingston, N. (1971). Selecting managers: A survey of current practice in 200 companies. *BIM, Management Survey Report*, No. 4.

Kinicki, A.J., & Lockwood, C.A. (1985). The interview process: An examination of factors recruiters use in evaluating job applicants. *Journal of Vocational Behaviour*, 26, 117-125.

Kinicki, A.J., Lockwood, C.A., Hom, P.W., & Griffeth, R.W. (1990). Interviewer predictions of applicant qualifications and interviewer validity: Aggregate and individual analyses. *Journal of Applied Psychology*, 75, 5, 477-486.

Klein, S.P. & Owens, W.A. (1965). Faking of a scored life history blank as a function of criterion objectivity. *Journal of Applied Psychology,* 49, 452-454.

Kleinman, M. (1993). Are Rating Dimensions in Assessment Centres Transparent for Participants? Consequences for Criterion and Construct Validity. *Journal of Applied Psychology*, 6, 988-993

Klimoski, R. J., & Brickner, M. (1987). Why do assessment centres work? The puzzle of assessment centre validity. *Personnel Psychology*, 40, 243-260.

Klimoski, R. J., & Strickland, W.J. (1977). Assessment centres - valid or merely prescient. *Personnel Psychology*, 30, 353-360.

Kline, P. (1992). *Psychometric testing in personnel selection and appraisal.* Surrey: Croner Publications Ltd.

Kluger, A.N., Reilly, R.R., & Russell, C.J. (1991). Faking biodata tests: Are option-keyed instruments more resistant? *Journal of Applied Psychology*, 76, 6, 889-896.

Konovsky, M.A., & Organ, D.W. (1996). Dispositional and contextual determinants of organisational citizenship behaviour. *Journal of Organizational Behaviour*, 17, 253-266.

Korman, A. K. (1968). The prediction of managerial performance: A review. *Personnel Psychology*, 21, 295-322.

Kotter, J.P. (1982). *The general managers.* New York: McGraw-Hill.

Kotter, J.P. (1988). *The leadership factor.* New York: Free Press.

Kotter, J.P. (1990). *Force for change.* New York: Free Press.

Kouzes, D., & Posner, B. (1987). *The leadership challenge: How to get extraordinary things done in organisations.* San Francisco: Jossey Bass.

Kraiger, K., & Ford, J.K. (1985). A meta-analysis of ratee race effects in performance ratings. *Journal of Applied Psychology,* 70, 56-65.

Kurke, L. B., & Aldrichs, H. E. (1983). Mintzberg was right! A replication and extension of the nature of managerial work. *Management Science*, 29, 8, 975-984.

Landy, F.J., & Farr, J.L. (1980). Performance testing. *Psychological Bulletin*, 87, 72-107.

Landy, F.J., Barnes, J.L., & Murphy, K. R. (1978). Correlates of perceived fairness and accuracy of performance evaluation. *Journal of Applied Psychology*, 63, 751-754.

Landy, F.J., Barnes-Farrell, J., & Cleveland, J.N. (1980). Perceived Fairness and accuracy of performance evaluation: A follow-up. *Journal of Applied Psychology*, 65, 3, 355-356.

Latham, G. P. & Saari, L. M. (1984). Do people do what they say? Further studies on the situational interview. *Journal of Applied Psychology,* 69, 569-573.

Latham, G. P., & Wexley, K.N. (1977). Behavioural observation scales for performance appraisal purposes. *Personnel Psychology*, 30, 255-268.

Latham, G. P., Saari, L. M., Pursell, E. D., & Campion, M. A. (1980). The situational interview. *Journal of Applied Psychology*, 65, 4, 422-427.

Latham, G.P., Wexley, K.N., & Pursell, E.D. (1975). Training managers to minimize rating errors in the observation of behaviour. *Journal of Applied Psychology*, 60, 550-555.

Lau, A.W., Newman, A. R., & Broedling, L.A. (1980). The nature of managerial work in the public sector. *Public Management Forum*, 19, 513-521.

Laurent, H. (1970). Cross-cultural cross-validation of empirically validated tests. *Journal of Applied Psychology*, 54, 5, 417-423.

Lautenschlager, G.J. & Atwater, D.C. (1986). Controlling reponse distortion on an empirically keyed biodata questionnaire. Unpublished manuscript. Cited in L. Hough & C. Paullin (1994). Construct-oriented scale construction: The rational approach. . In G.S. Stokes, M.D. Mumford,, and W.A. Owens (Eds.). *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction.* (pp. 109-145). Palo Alto, CA, USA: Cpp Books.

Lawler, E. E. (1967). The multitrait-multirater approach to measuring managerial job performance. *Journal of Applied Psychology*, 51, 5, 369-381.

Lawshe, C.H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28, 563-75.

Lent, R.H., Aurbach, H. D., & Levin, L. S. (1971). Predictors, criteria and significant results. *Personnel Psychology*, 24, 519-533.

Levin, H.M. (1988). Issues of agreement and contention in employment testing. *Journal of Vocational Behaviour*, 33, 398-403.

Levine, A.S. & Zachert, V. (1951). Use of a biographical inventory in the Air Force classification programme. *Journal of Applied Psychology*, 35, 241-244.

Liu, J., & Wilson, D. (2001). The unchanging perception of women as managers. *Women in Management Review*, 16, 4, 163-173.

Loevinger, J. (1994). Has psychology lost its conscience? *Journal of Personality Assessment*, 62, 1, 2-8.

Loevinger, J., Gleser, G.C., & Dubois, P.H. (1953). Maximizing the discriminating power of a multiple score set. *Psychometrika*, 18, 309-311.

London, M., & Stumpf, S. A. (1982). *Managing careers*. Reading, MA: Addison-Whiteley.

Lowery, C, M., & Krilowicz, T.J. 91994). Relationships among non-task behaviours, rated performance, and objective performance measures. *Psychological Reports*, 74, 571-578.

Mabe, P.A., & West, S.G. (1982). Validity of self-evaluation of ability: a review and meta-analysis. *Journal of Applied Psychology,* 67, 280-296.

Mael, F.A. (1991). A conceptual rationale for the domain and attributes of biodata. *Personnel Psychology,* 44, 736-792.

Mael, F.A.(1994). If past behaviour really predicts future, so should biodata's. In M. G. Rumsay, C. B. Walker, & Harris, J.H. (Eds.) *Personnel Selection and Classification*. Lawrence Erlbaum Associates Publishers, Hove, UK.

Mael, F.A., & Ashforth, B. E. (1995). Loyal from day one: Biodata, organisational identification, and turnover among newcomers. *Personnel Psychology*, 48, 309-333.

Mael, F.A., & Hirsch, A. C. (1993). Rainforest empiricism and quasi-rationality: Two approaches to objective biodata. Personnel Psychology, 46, 719-738.

Mael, F.A., Connerley, M., & Morath, R.A. (1996). None of your business: Parameters of biodata invasiveness. *Personnel Psychology*, 49, 613-650.

Mahoney, T. A., Jerdee, T .H., & Carroll, S.J., Jr. (1965). The jobs of management. *Industrial Relations*, 4, 97-110.

Mahoney, T. A., Jerdee, T. H., & Nash, A. N. (1960). Predicting managerial effectiveness. *Personnel Psychology*, 13, 147-163.

Makin, P.J. & Robertson, I. T. (1983). Self-assessment, realistic job previews and occupational decisions. *Personnel Review,* 12, 3, 21-25.

Martinko, M.J., & Gardner, W.L. (1990). Structured observation of managerial work: A replication and synthesis. *Journal of Management Studies*, 27, 3, 329-356.

Matteson, M.G., Osburn, H.G., & Sparks, C.P. (1969). A computer-based methodology for constructing homogenous keys with applications to biographical data (Report No 1). Houston, TX: University of Houston, personnel Psychology Services Centre. Cited in L.F. Schoenfeldt, & J.L. Mendoza (1994). Developing and using Factorially derived biographical scales. In G.S. Stokes, M.D. Mumford,, and W.A. Owens (Eds.). *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction.* (pp. 147-169). Palo Alto, CA, USA: Cpp Books.

Maxwell, S. E., & Arvey, R. D. (1993). The search for predictors with high validity and low adverse impact: Compatible or incompatible goals. *Journal of Applied Psychology,* 78, 3, 433-437.

May, K. (1998). Work in the 21st century: The role of I-O in work-life programs. Society for Industrial and Organisational Psychology, *The Industrial Psychologist*, 36, 2. Division of the American Psychological Association Organisational Affiliate of the American Psychological Society.

Mayer, J.D., Carlsmith, K.M., Frasier Chabot, H. (1998). Describing the person's external environment: Conceptualising and measuring the life space. *Journal of Research in Personality*, 32, 253-296.

McAdams, D. P. (1992). The five-factor model in personality: A critical appraisal. *Journal of Personality*, 60, 329-361.

McCall, M. W. Jr. (1994). Identifying leadership potential in future international executives: Developing a concept. *Consulting Psychology Journal*, 46, 49 - 63.

McCall, M.W., & Segrist, C.A. (1980). *In pursuit of the manager's job: Building on Mintzberg.* Greensboro, NC: Centre for Creative Leadership.

McCormick, E J. & Ilgen, D. (1987). *Industrial and Organizational Psychology.* 8th Edition. Prentice-Hall, Englewood Cliffs, NJ.

McCormick, E. J., & Ilgen, D. (1980). *Industrial Psychology.* 7th Ed. Prentice-Hall, Englewood Cliffs, NJ.

McCormick, E.J., Jeanerret, P. R., & Mecham, R.C. (1972). A study of job dimensions as based on the Position Analysis Questionnaire (PAQ). *Journal of Applied Psychology*, 56, 347-368.

McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., & Maurer, S.D. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology*, 79, 599-616.

McEvoy, G. M., & Beatty, R.W. (1989). Assessment centres and subordinate appraisals of managers: A seven-year examination of predictive validity. *Personnel Psychology*, 42, 1, 37-52.

McIntyre, R.M., Smith, D. E., & Hassett, C. E. (1984). Accuracy of performance ratings as affected by rater training and perceived purpose of rating. *Journal of Applied Psychology*, 69, 147-156.

McMannus, M.A., & Kelly, M.L. (1999). Personality measures and biodata: Evidence regarding their incremental predictive validity in the life industry. *Personnel Psychology*, 52, 137-148.

McMurray, R.N. (1947). Validating the patterned interview. *Personnel*, 23, 263-272.

Meyer, H.H. (1980). Self-appraisal of job performance. *Personnel Psychology*, 33, 291-295.

Meyer, H.H. (1987). Predicting supervisory ratings versus promotional progress in test validation studies. *Journal of Applied Psychology*, 72, 4, 696-697.

Miller, L. (1991). Managerial competencies. *Industrial and Commercial Training*, 23, 6, 11-15.

Mintzberg, H. (1975). The manager's job: Folklore and fact. *Harvard Business Review*, July-August, 49-61.

Mintzberg, H. (1973). *The nature of managerial work.* New York: Harper & Row.

Mintzberg, H. (1970). Structured observation as a method of studying managerial work. *Journal of Management Studies*, 7, 87-104.

Mischel, W. (1968). *Personality and assessment.* New York: Wiley.

Mischel, W. (1981). *Introduction to personality.* - 3rd ed. - New York; London : Holt, Rinehart and Winston.

Mitchel, J.O. (1975). Assessment centre validity: A longitudinal study. *Journal of Applied Psychology*, 60, 537-579.

Mitchell, J.L. (1978). Structured job analysis of professional and managerial positions. Unpublished doctoral dissertation, Purdue University, West Lafayette, IN. Cited in, Borman, W. C., & Brush, D. H. (1993). More progress toward a taxonomy of managerial performance requirements. *Human Performance*, 6, 1, 1-21.

Mitchell, T.W. & Klimoski, R.J. (1982). Is it rational to be empirical? A test of methods for scoring biographical data. *Journal of Applied Psychology*, 67, 411-418.

Mitchell, T.W. (1996). *Can do and will do criterion success: A practitioner's theory of biodata.* Paper presented at the Biennial Biodata Conference, Georgia, U.S.

Morrison, R.F., Owens, W.A., Glennon, J.R., & Albright, L.E. (1962). Factored life history antecedents of industrial research performance. *Journal of Applied Psychology*, 46, 281-284.

Mosel, J.N., & Goheen, H.W. (1958). The validity of the Employment Recommendation Questionnaire in personnel selection. I. Skilled trades. Personnel Psychology, 11, 481-490.

Mossholder, K. W., & Arvey, R. D. (1984). Synthetic validity: a conceptual and comparative review. *Journal of Applied Psychology*, 69, 322-333.

Motowidlo, S.J., & Vam Scotter, J.R. (1994). Evidence that task performance should be distinguished from contextual performance. *Journal of Applied Psychology*, 79, 4, 475-480.

Motowidlo, S.J., Carter, G.W., Dunnette, M.D., Tippens, N., Werner, S., Burnett, J. R., &Vaughan, M.J. (1992). Studies of the structured behaviour interview. *Journal of Applied Psychology*, 77, 571-587.

Mount, M.K., & Barrick, M.R. (1998). Five reasons why the 'big five' article has been frequently cited. *Personnel Psychology,* 51, 849-857.

Mount, M.K., Judge, T.A., Scullen, S.E., Sytsma, M.R., & Hezlett, S.A. (1998). Trait, rater and level effects in 360-degree performance ratings. *Personnel Psychology*, 51, 3, 557-577.

Mount, M.K., Witt, L.A., Barrick, M.R. (2000). Incremental validity of empirically keyed biodata scales over GMA and the five factor personality constructs. *Personnel Psychology*, 53, 299-323.

Muchinsky, P.M. (1979). The use of reference reports in personnel selection: a review and evaluation. *Journal of Occupational Psychology*, 52, 287-297.

Mumford, M. S., & Owens, W.A (1987). Methodology review: Principles, procedures, and findings in the application of background data measures. *Applied Psychological Measurement*, 11, 1, 1-31.

Mumford, M. S., Stokes, G. S., & Owens, W. A. (1990). *Patterns of life adaptations: The ecology of human individuality.* New Jersey, Erlbaum.

Mumford, M.D. & Stokes, G.S. (1992). Developmental determinants of individual action: Theory and practice in applying background measures. In M.D. Dunnette, and L.M. Hough (Eds.). *Handbook of industrial and organizational psychology*, Vol. 2 (2nd ed.). (pp. 61-138). Palo Alto, CA, USA: Consulting Psychologists Press, Inc.

Mumford, M.D. (1983). Social comparison theory and the evaluation of peer evaluations: a review and some applied implications. *Personnel Psychology*, 36, 867-881.

Mumford, M.D., Costanza, D. P., Connelly, M.S., Johnson, J.F. (1996). Item generation procedures and background data scales: Implications for construct and criterion-related validity. *Personnel Psychology*, 49, 361-398.

Mumford, M.D., Reiter-Palmon, R. & Snell, A. F. (1994). Background data and development: Structural issues in the application of life history measures. In G.S. Stokes, M.D. Mumford,, and W.A. Owens (Eds.). *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction.* (pp. 555-581). Palo Alto, CA, USA: Cpp Books

Munsterberg (1913). Psychology and Industrial Efficiency. Boston & New York: Houghton Mifflin Company. Cited in Schmitt, N., & Landy, F.J. (1993). The concept of validity. In N. Schmitt, W.C. Borman, and Associates. *Personnel Selection in Organizations.* (pp. 275-309). San Francisco: Jossey-Bass Publishers.

Murphy, R.T. (1994). Meta-analysis and validity generalization. In N. Anderson, and P. Herriot (Eds.). *Assessment and selection in organizations: Methods and practice for recruitment and appraisal.* First update and supplement 1994. John Wiley and Sons. Chichester:England.

Neidig, R.D., & Neidig, P.J. (1984). Multiple assessment centre exercises and job relatedness. *Journal of Applied Psychology*, 69, 1, 182-186.

Neter, E., & Ben-Shakhar, G. (1989). The predictive validity of graphological inference: A meta-analytic approach. *Personality and Individual Differences*, 10, 7, 737-745.

Neuman, G.A., & Kickul, J.R. (1998). Organisational citizenship behaviours: Achievement orientation and personality. *Journal of Business and Psychology*, 13, 2, 263-279.

New, G. A. (1996). Reflections: A three-tier model of organisational competencies. *Journal of Managerial Psychology*, 11, 8, 44-51.

Nisbett, R. E., & Wilson, W.C. (1977). On telling more than we can know. *Psychological Review*, 84, 231-256.

Noordegraaf, M., & Stewart, R. (2000). Managerial behaviour in private and public sectors: Distinctiveness, disputes and directions. *Journal of Management Studies*, 37, 3, 427-442.

Noordegraaf, M., & Stewart, R. (2000). Managerial behaviour in private and public sectors: Distinctiveness, disputes and directions. *Journal of Management Studies*, 37, 3, 427-442.

Okechuka, C. (1994). The relationship of six managerial characteristics to the assessment of managerial effectiveness in Canada, Hong Kong and People's Republic of China. *Journal of Occupational and Organisational Psychology*, 67, 79-86.

Ones, D.S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology*, 78, 4, 679-703.

Oppler, S.H., Campbell, J.P., Pulakos, E.D., & Borman, W.C. (1992). Three approaches to the investigation of subgroup bias in performance measurement: Review, results, and conclusions. *Journal of Applied Psychology*, 77, 201-217.

Organ, D.W. (1988), *Organizational Citizenship Behaviour: The Good Soldier Syndrome.* Lexington, MA: Lexington Books.

Organ, D.W. (1994). *Personality and organisational citizenship behaviour.* Journal of Management, 20, 2, 465-478.

Organ, D.W., & Konovsky, M. (1989). Cognitive versus affective determinants of organisational citizenship behaviour. *Journal of Applied Psychology*, 74, 1, 157-164.

Organ, D.W., & Lingl, A. (1995). Personality, satisfaction, and organisational citizenship behaviour. *The Journal of Social Psychology*, 135, 3, 339-350.

Organ, D.W., & Ryan, K. (1995). A meta-analytic review of attitudinal and dispositional predictors of organisational citizenship behaviour. *Personnel Psychology*, 48, 775-802.

Owens, W.A. (1968). Toward one discipline of scientific psychology. *American Psychologist*, 23, 782-785.

Owens, W.A. (1971). A quasi-actuarial prospect for individual assessment. *American Psychologist,* 26, 992-999.

Owens, W.A. (1976). Background data. In M.D. Dunnette (ED.) *Handbook of industrial and organisational psychology.* Chicago: Rand-McNally.

Owens, W.A., & Schoenfeldt, L.F. (1979). Toward a classification of persons. *Journal of Applied Psychology*, 64, 5, 569-607.

Owens, W.A., Glennon, J.R., & Albright, L. E. (1962). *A catalogue of life history items.* Greensboro, North Carolina, Richardson Foundation.

Pace, L.A., & Schoenfeldt, L.F. (1977). Legal concerns in the use of weighted applications. *Personnel Psychology*, 30, 2, 159-166.

Paulhaus, D. L. (1986). Self-deception and impression management in trait responses. In A. Angleitner, and J. S. Wiggins (Eds.), *Personality assessment via questionnaire.* New York: Springer.

Pearlman, K., Schmidt, F.L., & Hunter, J.E. (1980). Validity generalization results for tests used to predict job proficiency and training success in clerical occupations. *Journal of Applied Psychology*, 65, 373-406.

Penfield, R.V. (1975). Time allocation patterns and effectiveness of managers. *Personnel Psychology*, 27, 245-255.

Pettigrew, A.M. (1973). *The politics of organisational decision-making.* London:

Pfeffer, J., & Salancik, G.R. (1978). *External control of organisations: A resource dependence perspective.* New York: Harper and Row.

Phares, E. & Chaplin, J. (1997). *Introduction to personality.* Fourth edition. New York, Longman.

Prieto, J-M. (1989). Tests of Aptitude. In P. Herriot (Ed.) *Assessment and selection in organisations: Methods and practice for recruitment and appraisal.* Chichester, England: John Wiley and Sons Ltd.

Pulakos, E. D. (1984). A comparison of rater training programs: Error training and accuracy training. *Journal of Applied Psychology*, 69, 581-588.

Pulakos, E.D., White, L. A., Oppler, S.H., & Borman, W.C. (1989). Examination of race and sex effects on performance ratings. *Journal of Applied Psychology*, 74, 70-780.

Pye, A. (1991). Management competence: 'The flower in the mirror and the moon on the water'. In M. Silver (Ed.). *Competent to manage.* Routledge: UK.

Raven, J.L. (1960). *Guide to the Standard Progressive Matrices.* London: H.K. Lewis.

Ree, M.J., & Earles, J. A. (1994). The ubiquitous predictiveness of g. In: M.G. Rumsey, C. B. Walker, & J.H. Harris (Eds.) *Personnel selection and classification* (pp. 127-135). New Jersey: Lawrence Erlbaum Associates.

Reed, M. I. (1984). Management as social practice. *Journal of Management Studies*, 21, 3, 273-286.

Reed, M. I., & Anthony, P. (1992). Professionalizing management and managing professionalization: British management in the 1980s. *Journal of Management Studies*, 29, 5, 301-323.

Rees, W. D. (1996). *The skills of management* (4th ed.). International Thomson Business Press, London, UK.

Reilly, R.R., & Chao, G.T. (1982). Validity and fairness of some alternative employee procedures. *Personnel Psychology*, 35, 1-62.

Reilly, R.R., Henry, S., & Smither, J.W. (1990). An examination of the effects of using behaviour checklists on the construct validity of assessment centre dimensions. *Personnel Psychology*, 43, 1, 71-84.

Reiter-Palmon, R., & Connelly, M.S. (2000). Item selection counts: A comparison of empirical key and rational scale validities in theory-based and non-theory based item pools. *Journal of Applied Psychology*, 85, 1, 143-151.

Ridgeway, C.C. (1992). The competence profile of international senior managers/directors. *Recruitment, Selection, & Retention*, 1, 3, 22-23.

Rim, Y. (1961). Dimensions of job incentives and personality. *Acta Psychologica*, 18, 332-336.

Ritchie, R.R., & Boehm, V.R. (1977). Biographical data as a predictor of women's and men's management potential. *Journal of Vocational Behaviour*, 11, 363-368.

Ritchie, R. J., & Moses, J.L. (1983). Assessment centre correlates of women's advancement into middle management: A 7-year longitudinal analysis. *Journal of Applied Psychology*, 68,2, 227-231

Roadman, H.E. (1964). An industrial use of peer ratings. *Journal of Applied Psychology*, 48, 211-214.

Robertson, I. T., Baron, H., Gibbons, P., MacIver, R., & Nyfield, G. (2000). Conscientiousness and managerial performance. *Journal of Occupational and Organizational Psychology*, 73, 171-180.

Robertson, I.T., & Iles, P.A. (1988). Approaches to Managerial Selection. In C. L. Cooper, D., & Robertson, I., (Eds.) *International Review of Industrial and Organizational Psychology.* London: John Wiley & Sons Ltd.

Robertson, I.T., & Kandola, R. S. (1982). Work sample tests of trainability: a meta-analysis. *Journal of Occupational Psychology*, 55, 171-183.

Robertson, I.T., & Kinder, A. (1993). Personality and job competences: The criterion-related validity of some personality variables. *Journal of Occupational and Organisational Psychology*, 66, 225-244.

Robertson, I.T., & Makin, P.J. (1986). Management Selection in Britain: A Survey and Critique. *Journal of Occupational Psychology*, 59, 45-57.

Robertson, I.T., & Makin, P.J. (1993). Selection Methods and Their Usage. *Recruitment, Selection & Retention*, 2, 1, 3-10.

Robertson, I.T., Baron, H., Gibbons, P., MacIver, R., & Nyfield, G. (2000). Conscientiousness and managerial performance. *Journal of Occupational and Organisational Psychology*, 73, 2, 171-180.

Rose, D.S., & Baydoun, R. (1995). Content validity: A neglected strategy for developing managerial selection tests? *Current Psychology: Research and Reviews*, 14, 2, 138-151.

Rosen, B., & Jerdee, T.H. (1973). The influence of sex stereotypes on evaluations of male and female supervisory behaviour. *Journal of Applied Psychology*, 57, 44-48.

Rosen, M. (1984). Myth and reproduction: The contextualization of management theory, method and practice. *Journal of Management Studies*, 21, 3, 303-322.

Rothaus, P., Morton, R.B., & Hanson, P.G. (1965). Performance appraisal and psychological distance. *Journal of Applied Psychology*, 49, 48-54.

Rothstein, H.R., Schmidt, F.L., Erwin, F.W., Owens, W.A., & Sparks, C.P. (1990). Biographical data in employment selection: can validities be made generalisable? *Journal of Applied Psychology*, 75, 175-184.

Russell, C. J. (1983). *An examination of assessor's clinical judgements*. Presented at the 11th International Congress on the Assessment Centre Method, Williamsburg, VA.

Russell, C. J. (1985). Individual decision processes in an assessment centre. *Journal of Applied Psychology*, 70, 737-746.

Russell, C. J., & Domm, D. R. (1990). *On the validity of role-congruency-based assessment centre procedures for predicting performance appraisal ratings, sales revenue,* and profit. Presented at the 50th annual meeting of the Academy of Management, San Francisco, CA.

Russell, C.J. (1994). Generation procedures for biodata items: A point of departure. In G.S. Stokes, M.D. Mumford, & W.A. Owens (Eds.). *The Biodata Handbook: Theory, research and use of biographical information in selection and performance prediction*. (pp. 17-38). Palo Alto, CA, USA: Cpp Books.

Ryan, A.M., McFarland, L., Baron, H., & Page, R. (1999). An international look at selection practices: Nation and culture as explanations for variability in practice. *Personnel Psychology*, 52, 359-391.

Saal, F.E., & Landy, F.J. (1977). The mixed standard rating scale: An evaluation. *Organizational Behaviour and Human Performance*, 18, 413-414.

Sackett, P. R., & Dreher, G.F. (1982). Constructs and assessment centre dimensions: some troubling empirical findings. *Journal of Applied Psychology*, 67, 401-410.

Sackett, P. R., & DuBois, C.L.Z. (1991). Rater-ratee race effects on performance evaluation: Challenging meta-analytic conclusions. *Journal of Applied Psychology*, 76, 6, 873-877.

Sackett, P. R., & Hakel, M.D. (1979). Temporal stability and individual differences in using assessment information to form overall ratings. *Organizational Behaviour and Human Performance*, 23, 120-137.

Sackett, P.R., Harris, M.M., & Orr, J.M. (1986). On seeking moderator variables in the meta-analysis of correlational data: A monte carlo investigation of statistical power and resistance to Type I error. *Journal of Applied Psychology*, 71, 2, 302-310.

Salgado, J.F. (1996). Personality and job competences: A comment on the Robertson and Kinder (1993) study. *Journal of Occupational and Organisational Psychology*, 69, 373-375.

Satow, R., & Rector, J. (1995). Using Gestalt Graphology to identify Entrepreneurial Leadership. *Perceptual and Motor Skills*, 81, (1) 263-270.

Saville and Holdsworth Limited (1985). *Occupational Personality Questionnaire: Manual.* SHL Limited, Esher, Surrey.

Sawyer, J. (1966). Measurement and prediction, clinical and statistical. *Psychological Bulletin*, 66, 178-200.

Scarborough, H. (1998). The unmaking of management? Change and continuity of British management in the 1990s. *Human Relations,* 51, 6, 691-717.

Schmidt, F.L. & Hunter, J.E (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research. *Psychological Bulletin*, 124, 2, 262-274.

Schmidt, F. L., Hunter, J.E., & Outerbridge, A.N. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology*, 71, 432-439.

Schmidt, F. L., Hunter, J.E., & Pearlman, K. (1981). Task differences as moderators of aptitude test validity in selection: a red herring. *Journal of Applied Psychology*, 66, 166-185.

Schmidt, F. L., Hunter, J.E., & Raju, N. S. (1988). Validity generalization and meta-analysis with commentaries. *Journal of Applied Psychology*, 73, 665-672.

Schmidt, F.L., Law, K., Hunter, J.E., Rothstein, H.R., Pearlman, K., & McDaniel, M. (1993). Refinements in validity generalisation methods: Implications for the situational specificity hypothesis. *Journal of Applied Psychology*, 78, 1, 3-12.

Schmidt, F.L., & Rothstein, H.R. (1994). Application of validity generalization to biodata scales in employment selection. In G.S. Stokes, M.D. Mumford,, and W.A. Owens (Eds.). *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction.* (pp. 237-260). Palo Alto, CA, USA: Cpp Books.

Schmitt, M.J., & Ryan, A.M. (1993). The big five in personnel selection: Factor structure in applicant and non-applicant populations. *Journal of Applied Psychology*, 78, 6, 966-974.

Schmitt, N. (1977). Interrater agreement in dimensionality and combination of assessment centre judgements. *Journal of Applied Psychology*, 62, 171-176.

Schmitt, N., & Hill, T.E. (1977). Sex and race composition of assessment centre groups as a determinant of peer and assessor ratings. *Journal of Applied Psychology*, 62, 261-164.

Schmitt, N., & Landy, F.J. (1993). The concept of validity. In N. Schmitt, W.C. Borman, and Associates. *Personnel Selection in Organisations.* (pp. 275-309). San Francisco: Jossey-Bass Publishers.

Schmitt, N., & Lappin, M. (1980). Race and sex as determinants of the mean and variance of performance ratings. *Journal of Applied Psychology*, 65, 4, 428-435.

Schmitt, N., & Noe, R.A. (1986). Personnel selection and equal employment opportunity. In C. L. Cooper and I.T. Robertson, (Eds.). *International Review of Industrial and Organisational Psychology.* Chichester: Wiley.

Schmitt, N., Gooding, R.Z., Noe, R.A., & Kirsch, M. (1984). Metaanalysis of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology*, 37, 407-422.

Schmitt, N., Noe, R.A., & Gottschalk, R. (1986). Using the lens model to magnify raters' consistency, matching, and shared bias. *Academy of Management Journal*, 29, 130-139.

Schmitt, N., Schneider, J. R., & Cohen, S. A. (1990). Factors affecting validity of a regionally administered assessment centre. *Personnel Psychology*, 43, 1, 1-12.

Schnake, M., Dimler, M.P., & Cochran, D.S. (1993). The relationship between 'traditional' leadership, 'super' leadership and organisational citizenship behaviour. *Group and Organisation Management*, 18, 3, 352-365.

Schneier, C. E. (1977). Operational utility and psychometric characteristics of behavioural expectation scales: A cognitive reinterpretation. *Journal of Applied Psychology*, 62, 541-548.

Schoenfeldt, L.F.(1989). Biographical data as the new frontier in employee selection research. Paper presented at the annual meeting of thw American Psychological Association, new Orleans. Cited in M.D., Mumford, Reiter-Palmon, R. & Snell, A. F. (1994). Background data and development: Structural issues in the application of life history measures. In G.S. Stokes, M.D. Mumford,, and W.A. Owens (Eds.). *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction.* (pp. 555-581). Palo Alto, CA, USA: Cpp Books

Schrader, A.D., & Osburn, H.G. (1977). Biodata faking: Effects of induced subtlety and position specificity. *Personnel Psychology*, 30, 3, 395-404.

Schrader, B.W., & Steiner, D.D. (1996). Common comparison standards: An approach to improving agreement between self and supervisory performance ratings. *Journal of Applied Psychology*, 81, 6, 813-820.

Schwab, D.P., & Heneman, H.G.III (1978). Age stereotyping in performance appraisal. *Journal of Applied Psychology*, 63, 573-578.

Scriven, R., Savage, A., & McCredie, H. (1994). The search for 'executive intelligence'. *Selection and Development Review*, 10, 2, 2-4.

Shackleton, V., & Newell, S. (1991). Management Selection: A comparative survey of methods used in top British and French companies. *Journal of Occupational Psychology*, 64, 23-36.

Shackleton, V., & Newell, S. (1989). Selection procedures in practice. In P. Herriot (Ed.). *Assessment and Selection in organisations: Methods and Practice for recruitment and appraisal* (pp. 245-256). John Wiley and sons. Chichester: England.

Shapira, Z., & Shirom, A. (1980). New issues in the use of behaviourally anchored rating scales: Level of analysis, the effects of incident frequency and external validation. *Journal of Applied Psychology*, 65, 517-523.

Sharf, J.C. (1994). The impact of legal and equal opportunity issues on personal history inquiries. In G.S. Stokes, M.D. Mumford, & W.A. Owens (Eds.), *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction* (pp. 351-390). Palo Alto, CA: CPP Books.

Sharples, H.D. (1992). The management paradigm for the 1990s. *Recruitment, Selection & Retention,* 1, 2, 13-20.

Silver, E.M., & Bennett, C. (1987). Modification of the Minnesota Clerical Test to predict performance on video display terminals. *Journal of Applied Psychology*, 72, 153-155.

Singer, M., Singer, A., & Bruhns, C. (1991). Fair managerial selection. *Journal of Managerial Psychology*, 6, 5, 22-25.

Smith, D., & Tarpey, T. (1987). In-tray excercies and assessment centres: The issue of reliablity. *Personnel Review*, 16, 3, 24-28.

Smith, M. (1991). Recruitment and Selection in the UK with some data on Norway. *European Review of Applied Psychology*, 41, (1), 27-34.

Smith, M. (1994). A theory of the validity of predictors in selection. *Journal of Occupational and Organisational Psychology*, 67, 13-31.

Smith, M., & Robertson, I.T. (1986). *The theory and practice of systematic staff selection.* The Macmillan Press Ltd: Hampshire.

Smith, M., Gregg, M., & Andrews, D. (1989). *Selection and assessment: A new appraisal.* London: Pitman Publishing.

Smith, P.C., & Kendall, L.M. (1963). Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology*, 47, 149-155.

Smith, W.J., Albright, L.E., Glennon, J.R., & Owens, W.A. (1961). The prediction of research competence and creativity from personal history. *Journal of Applied Psychology*, 45, 59-62.

Sneath, F., Thakur, M., & Medjuck, B. (1976). *Testing people at work.* London: Institute of Personnel Management.

Sorcher, M. (1985). *Predicting executive success.* New York: Wiley.

Sparrow, J., Patrick, J., Spugeon, P., & Barwell, F. (1982). The use of job component analysis and related aptitudes in personnel selection. *Journal of Occupational Psychology*, 55, 157-164.

Spreitzer, G. M., McCall, M.W., & Mahoney, J. D. (1997). Early identification of international executive potential. *Journal of Applied Psychology*, 82, 1, 6-29.

Spreitzer, G.M., Quinn, R. E., & Fletcher, J. (1995). Excavating the paths of meaning, renewal, and empowerment: A typology of managerial high-performance myths. *Journal of Management Inquiry,* 4, 1, 16-39.

Springer, D. (1953). Ratings of candidates for promotion by co-workers and supervisors. *Journal of Applied Psychology*, 37, 347-351.

Stagner, R. (1969). Corporate decision making. *Journal of Applied Psychology*, 53, 1-13.

Stahl, M.J. (1983). Achievement, power and managerial motivation: Selecting managerial talent with the job choice exercise. *Personnel Psychology*, 36, 775-789.

Steiner, D. D., & Gilliland, S. W. (1996). Fairness reactions to Personnel Selection techniques in France and the United States. *Journal of Applied Psychology*, 81, 2, 134-141.

Steiner, D. D., Rain, J.S., & Smalley, M.M. (1993). Distributional ratings of performance: Further examination of a new rating format. *Journal of Applied Psychology,* 78, 3, 438-442.

Stephens, C. S. (1993). Five CIO's at work: Folklore and facts revisited. *Journal of Systems Management*, 44, 34-42.

Stewart, R. (1967). *Managers and their jobs: A study of the similarities and differences in the way managers spend their time.* London: MacMillan.

Stewart, R. (1982). A model for understanding managerial jobs and behaviour. *Academy of Management Review*, 7, 7-14.

Stewart, R. (1982). *Choices for the manager: A guide to understanding managerial work and behaviour.* Englewood Cliffs, NJ: Prentice Hall.

Stewart, R. (1989). Studies of managerial jobs and behaviour: The ways forwards. *Journal of Management Studies*, 26, 1, 1-10.

Stiles, P., Gratton, L., Truss, C., Hope-Haley, V., & McGovern, P. (1997). Performance management and the psychological contract. *Human Resource Management Journal*, 7, 1, 57-66.

Stokes, G.S., & Cooper, L.A. (1994). Selection using biodata: Old notions revisited. In G.S. Stokes, M.D. Mumford, and W.A. Owens. *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction*. (pp. 311-349). Palo Alto, CA, USA: Cpp Books.

Stricker, L.J., & Rock, D. A. (1998). Assessing leadership potential with a biographical measure of personality traits. *International Journal of Selection and Assessment*, 6, 3, 164-184.

Sulsky, L.M., & Balzer, W.K. (1988). Meaning and measurement of performance rating accuracy: Some methodological and theoretical concerns. Journal of Applied Psychology, 73, 497-506.

Tajfel, H. (1978). *Differentiation between social groups: studies in the social psychology of intergroup relations*. London: Published in cooperation with European Association of Experimental Social Psychology by Academic Press.

Taylor, F. W. (1911). *The principles of scientific management.* New York: Harper & Rowe.

Taylor, P., Mills, A., & O'Driscoll, M. (1993). Personnel selection methods used by New Zealand organisations and personnel consulting firms. *New Zealand Journal of Psychology*, 22, 19-31.

Tenopyr, M. L. (1977). Content-construct confusion. *Personnel Psychology*, 30, 47-54.

Tett, R. P. (1998). Is conscientiousness ALWAYS positively related to job performance? Society for Industrial and Organisational Psychology, The Indutrial Psychologist (TIP), 36, 2. Division of the American Psychological Association Organisational Affiliate of the American Psychological Society. http://www.siop.org/tip/backissues/TIPJuly98/tett.htm

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: a meta-analytic review. *Personnel Psychology*, 44, 703-742.

Thayer, P.W. (1977). Something old, something new. *Personnel Psychology*, 30, 513-524.

The IPD Guide on Recruitment. (1996). The Institute of Personnel and Development. London: IPD.

The Tenth Mental Measurements Yearbook. Lincoln: Buros Institute of Mental Measurements of the University of Nebraska.

Thornton, G.C. (1980). Psychometric properties of self-appraisals of job performance. *Personnel Psychology*, 33, 263-271.

Thornton, G.C., III, & Byham, W.C. (1982). *Assessment centres and managerial performance.* New York: Academic Press.

Tornow, W.W., & Pinto, P. R. (1976). The development of a managerial job taxonomy: A system for describing, classifying and evaluating executive positions. *Journal of Applied Psychology*, 61, 410-418.

Torrington, D., & Weightman, J. (1994). *Effective management: People and organisations* (Second edition). Prentice Hall, UK.

Tosi, H.L., & Einbender, S. W. (1985). The effects of the type and the amount of information in sex discrimination research: a meta-analysis. *Academy of Management Journal,* 28, 712-723.

Townley, B. (1993). Performance appraisal and the emergence of management. *Journal of Management Studies,* 30, 2, 221-238.

Tucker, D. H., & Rowe, P.M. (1979). Relationships between expectancy, causal attributions, and final hiring decisions in the employment interview. *Journal of Applied Psychology,* 64, 27-34.

Turnage, J.J., & Muchinsky, P.M. (1984). A comparison of the predictive validity of assessment centre evaluations versus traditional measures in forecasting supervisory job performance: interpretative implications of criterion distortion for the assessment paradigm. *Journal of Applied Psychology*, 69, 595-602.

Turner, J. (1987). *Rediscovering the social group: a self-categorization theory.* Oxford, UK: Basil Blackwell

Turner, J. (1991). *Social influence.* Milton Keynes: Open University Press.

Tziner, A., Meir, E.I., Dahan, M., & Birati, A. (1994). An investigation of the predictive validity and economic utility of an assessment centre for the high-management level. *Canadian Journal of Behavioural Science*, 26, 2, 228-245.

Ulrich, I., & Trumbo, D. (1965). The selection interview since 1949. *Psychological Bulletin*, 63, 100-116.

Ungerson, B. (Ed.) (1975). *Recruitment Handbook.* Gower Press, Epping.

Urwick, L.F. (1952). *Notes on the theory of organisation.* New York: American Management Association.

Vale, C.D., Keller, L. S., & Bentz, V.J. (1986). Development and validation of a computerised interpretation system for personnel tests. *Personnel Psychology*, 39, 525-542.

Van Der Maesen De Sombreff, P., & Hofstee, W.K.B. (1989). Personality questionnaire and inventories. In P. Herriot (Ed.) *Assessment and selection in organisations: Methods and practice for recruitment and appraisal.* Chichester, England: John Wiley and Sons Ltd.

Van Dyre, L., Graham, J.W., & Dienesch, R.M. (1994). Organisational citizenship behaviour: Construct redefinition, measurement and validation. *Academy of Management Journal*, 37, 4, 765-802.

Van Rijn, P. (1980). Biographical questionnaires and scored application blanks in personnel selection. (PRR-80-31). Washington: US Office of personnel Management. Cited in J.A. Gandy, D.A. Dye, & C.N. MacLane. (1994). Federal government selection: The individual achievement record. In G.S. Stokes, M.D. Mumford,, and W.A. Owens (Eds.). *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction.* (pp. 275-309). Palo Alto, CA, USA: Cpp Books

Vance, R. J., MacCallum, R.C., Coovert, M.D., & Hedge, J.W. (1988). Construct validity of multiple job performance measures using confirmatory factor analysis. *Journal of Applied Psychology,* 73, 1, 74-80.

Vernon, P. E. (1959). The validation of Civil service Selection Board procedures. *Occupational Psychology,* 24, 75-95.

Vinchur,A. J., Schippman, J.S., Switzer, F.S., & Roth, P.L. (1998). A meta-analytic review of predictors for salespeople. *Journal of Applied Psychology*, 83, 4, 586-597.

Viswesvaran, C., Ones, D.S., & Schmidt, F. L. (1996). Comparative analyses of the reliability of job performance ratings. *Journal of Applied Psychology*, 81, 5, 557-574.

Wagner, S.H., & Goffin, R. D. (1997). Differences in accuracy of absolute and comparative performance appraisal methods. *Organisational Behaviour and Human Decision Processes*, 70, 2, 95 - 103.

Waldman, D. A., Atwater, L.E., & Antonioni, D. (1998). Has 360-degree feedback gone amok? *The Academy of Management Executive*, 12, 2, 86-85.

Walsh, J.P., Weinberg, R.M., & Fairfield, M. L. (1987). The effects of gender on assessment centre evaluations. *Journal of Occupational Psychology*, 60, 305-309.

Wanous, J.P. (1989). Installing a realistic job preview: Ten tough choices. *Personnel Psychology,* 42, 117-134.

Watson & Glaser (1980) *The Watson Glaser Critical Thinking Appraisal Test Manual.* The Psychological Corporation, London.

Watson, T.J. (1994). *In search of management: Culture, chaos and control in managerial work.* Routledge, London, UK.

Wayne, S.J., & Liden, R.C. (1995). Effects of impression management on performance ratings: A longitudinal study. *Academy of Management Journal*, 38, 1, 232-260.

Webster, E. C. (1964). *Decision making in the employment interview.* Montreal: Eagle.

Wechsler, D. (1955). *Wechsler Adult Intelligence Scale.* New York: Psychological Corporation.

Weisner, W.H., & Cronshaw, S.F. (1988). A meta-analytic investigation of the impact of interview format and the degree of structure on the validity of the employment interview. *Journal of Psychology*, 61, 275-290.

Werner, J.M. (1994). Dimensions that make a difference: Examining the impact of in-role and extra-role behaviours on supervisor ratings. *Journal of Applied Psychology*, 79, 1, 98-107.

Wernimont, P.F., & Campbell, J.P. (1968). Signs, Samples and Criteria. *Journal of Applied Psychology*, 52, 372-376.

Wherry, R. J., & Naylor, J.C. (1966). Comparison of two approaches - JAN and PROF-for capturing rater strategies. *Educational and Psychological Measurement*, 26, 267-286.

Whitely, W. (1985). Managerial work behaviour: An integration of results from two major approaches. *Academy of Management Journal*, 28, 2, 344-362.

Wilkinson, L.J. (1993). An alternative view of biodata: The model. *Recruitment, Selection and Retention,* 3, 1, 11-16.

Wilkinson, L.J. (1994). An alternative model of biodata: Some empirical evidence. *Recruitment, Selection and Retention,* 2, 3, 23-28.

Wilkinson, L.J. (1995). *The development of an alternative biodata model for managerial selection.* University of Newcastle Upon Tyne, Unpublished Doctoral Thesis.

Wilkinson, L.J. (1997). Generalisable biodata? An application to the vocational interests of managers. *Journal of Occupational and Organisational Psychology,* 70, 49-60.

Wilkinson, L.J., & van Zwanenberg, N. (1994). Development of a person specification system for managerial jobs. *Personnel Review*, 23, 1, 25-36.

Williams, R. E. (1956). A description of some executive abilities by means of the critical incident technique. Unpublished doctoral dissertation, Columbia University, New York. Cited in Borman, W.C., & Brush, D. H. (1993). More progress toward a taxonomy of managerial performance requirements. *Human Performance*, 6, 1, 1-21.

Willmott, H. C. (1984). Images and ideals of managerial work: A critical examination of conceptual and empirical accounts. *Journal of Management Studies*, 21, 3, 349-368.

Willmott, H. C. (1987). Studying managerial work. *Journal of Management Studies*, 24, 3, 249 -270.

Wilson, D. G., & Wagner, E. E.(1981).The Watson-Glaser critical thinking appraisal as a predictor of performance in a critical thinking course. *Educational and Psychological Measurement*, 41,1319-22

Wilson, N. A. B. (1948). The work of the Civil service Selection Board. *Occupational Psychology*, 22, 204-212.

Woehr, D. J., & Huffcutt, A. I. (1994). Rater training for performance appraisal: A quantitative review. *Journal of Occupational and Organisational Psychology*, 67, 189-205.

Wohlers, A. J., & London, M. (1989). Ratings of managerial characteristics: Evaluation difficulty, co-worker agreement, and self-awareness. Personnel Psychology, 42, 235-261.

Wollowick, H. B., & McNamara, W. J. (1969). Relationship of components of an assessment centre to management success. *Journal of Applied Psychology*, 53, 348-352.

Wonderlic, E.F. (1959). *Wonderlic Personnel Test.* Illinois: E.F. Wonderlic.

Woodruffe, C. (1993). What is meant by a competency? *Leadership and Organisation Development Journal*, 14, 29-36.

Yammarino, F.J. & Atwater, L. E. (1997). Do managers see themselves as others see them? Implications for self-other rating agreement for human resource management. *Organizational Dynamics*, 25, 4, 35-45.

Yammarino, F.J., & Atwater, L.E. (1997). Self-other rating agreement: A review and model. In G.R. Ferris, & (eds.). *Research in personnel and human resources management,* Vol. 15. (pp. 121-174). Greenwich, CT, USA: Jai Press, Inc. ix, 383 pp

Yammarino, F.J., & Waldman, D. A. (1993). Performance in relation to job skill importance: A consideration of rater source. *Journal of Applied Psychology*, 78, 2, 242-249.

Yukl, G. A. (1983, October). A taxonomy for integrating diverse perspectives on managerial behaviour. Paper presented at the annual meeting of the American Psychological Association, New York. Cited in, W. C. Borman & D. H. Brush, (1993). More progress toward a taxonomy of managerial performance requirements. *Human Performance*, 6, 1, 1-21.

Zedeck, S., & Cascio, W.F. (1982). Performance appraisal decisions as a function of rater training and purpose of the appraisal. *Journal of Applied Psychology*, 67, 6, 752-758.

Zedeck, S., & Kafry, D. (1977). Capturing rater policies for processing evaluation data. *Organizational Behaviour and Human Performance*, 18, 269-294.

Zedeck, S., Imparato, N., Krausz, M., & Oleno, T. (1974). Development of behaviourally anchored rating scales as a function of organizational level. *Journal of Applied Psychology,* 59, 249-252.

Zuckerman, M. (1989). Review of the 16PF Questionnaire. In J.C. Connoley and J.J. Kramer (Eds.), *The Tenth Mental Measurements Yearbook.* Lincoln: Buros Institute of Mental Measurement, NY.

Appendix I

Items added to the Wilkinson (1995) biodata form

59.	Do you worry about mistakes you have made?	1 2 3	Very often Often Occasionally
60.	Most of the time I am bursting with energy.	1 2	No Yes
62.	How important is job security to you?	1 2 3	Very important Important Unimportant
64.	Are you-	1 2 3	Very ambitious Ambitious Comfortable as you are
65.	How important is it that you get recognition for success in your field?	1 2 3 4	Very important Important Not so important Not important at all
66.	Do you consider yourself to be-	1 2 3 4	More of a risk taker than most As much of a risk taker as most Less of a risk taker than most Not a risk taker

Appendix II

Pilot Study Biodata Questionnaire

Candidate Code:

Please make one response per question unless otherwise indicated.

This first section is collecting general background information.

1.	Please indicate your sex	1	Female
		2	Male
2.	Indicate your age group	1	Under 18
		2	18 - 19
		3	20 - 24
		4	25 - 29
		5	30 - 39
		6	40 - 49
		7	over 50
3.	What was your family status in your parental home?	1	Only child
		2	Oldest of several
		3	Youngest of several
		4	Middle child
4.	Indicate the nature of the main wage earner's occupation in your parental home.	1	Professional
		2	Skilled
		3	Semi skilled
		4	Unskilled
5.	Are you currently a full time student?	1	No
		2	Yes
6.	Are you left handed or right handed?	1	Left handed
		2	Right handed

The next section is collecting information about your educational background.

7.	When at school, which subject were you most interested in?	1	Physics
		2	Chemistry or biology
		3	Technology or design
		4	Computing or IT
		5	Mathematics
		6	History
		7	Languages
8.	Please indicate your main reasons for continuing your education ***(mark all that apply)***	1	Career prospects
		2	Challenge and responsibility
		3	Job security
		4	Interesting and varied work
		5	To make best use of my skills
		6	Like creating things
		7	Like solving practical problems

9.	How many 'GCSE' passes, at grade C or better, or equivalents (Scottish 'O' grades) do you have?	1 2 3 4	Less than 3 3 - 6 7 - 10 More than 10

Please answer just one of questions 10a or 10b

10a.	How many 'A' level passes do you have?	1 2 3 4 5	0 1 2 3 4 or more
10b.	How many Scottish Highers do you have?	1 2 3 4 5	0 1 - 2 3 - 4 5 - 6 7 or more

10c.	Do you hold-	1 2 3 4	GNVQ National Certificate or Diploma Higher National Cert. or Diploma Other e.g. International Baccalaureate

11.	At what type of institution did you study for your University entry qualifications (e.g. 'A' levels etc.)?	1 2 3 4 5	Not applicable Secondary School VI Form College Technical or FE College Polytechnic or University

12.	Indicate where your pre-university education took place.	1 2 3 4 5	Scotland England, Wales, Northern Ireland European Community (excl. UK) Europe (excl. European Comm.) Outside Europe
13.	How do you normally revise for your examinations?	1 2 3 4	Plan to cover everything Concentrate on the weak areas Prepare selected topics Cover as much as you can in the time available.
14.	At which of the following have you completed a full course of study? ***(mark all that apply)***	1 2 3 4	University Polytechnic College of HE None of these

15.	Indicate the nature of your highest formal educational qualification.	1	Doctorate
		2	Masters/PG Dip
		3	Degree
		4	Diploma/Certificate
		5	A levels or equivalents
		6	University foundation course
16.	Have you ever studied for pleasure at part-time or evening classes ?	1	No
		2	Yes
17.	Have you ever studied statistics?	1	No
		2	Yes

This section is concerned with how you work on tasks

18.	Which of the following best describes your work style?	1	Work on one task at a time
		2	Work on several tasks at once
		3	Work on many tasks at once
19.	In a work situation I prefer to-	1	Work independently
		2	Work within a loose framework
		3	Work in a well defined framework
20.	Which of the following best describes your approach to work scheduling?	1	Tackle tasks as they arise
		2	Tackle tasks in order of priority
		3	Tackle the easy or short tasks first.
21.	What is the size of the project group that you work best with?	1	Large team
		2	Small team
		3	Group of 2 or 3 people
		4	Individual
22.	In which of the following media do you prefer to work?	1	Words
		2	Pictures
		3	Numbers/symbols
		4	No preference
23.	Which of the following describes your approach to an unfamiliar problem?	1	Guess a solution and check it works activities best
		2	Eliminate non-feasible solutions
		3	Attempt to deduce the solution from the information given
		4	Attempt to relate to a similar known problem

24.	When working on a problem do you prefer to-	1	Break it up into small stages, solving each stage in turn
		2	Attempt to solve the problem in its entirety
25.	When working on a project I prefer the situation to-	1	Be very well specified and predictable
		2	Have a degree of predictability
		3	Have a degree of uncertainty
		4	Be uncertain and unpredictable
26.	When working on a problem, I like to -	1	Tackle it in logical stages
		2	Focus on specific aspects that interests me
		3	Explore a range of different approaches
		4	Sit back and think of the best way of approaching it
		5	Come to an answer intuitively
27.	When things get really difficult I-	1	Carry on regardless
		2	Try to find a way around the problem
		3	Move onto something else
28.	When acquiring a new work skill do you practice until-	1	Nearly perfect
		2	Adequate
		3	Bored with practice
		4	Something else- please specify
29.	When learning to use an unfamiliar piece of complicated equipment do you-	1	Read the instruction manual
		2	Experiment with the controls
		3	Get someone to demonstrate it
		4	Other - please specify
30.	Which of the following best describes your needs when undertaking a new activity?	1	Careful supervision
		2	Some supervision
		3	Minimal supervision
		4	Other - please specify
31.	What is your strategy when working on a long and demanding task?	1	Work at it until complete
		2	Work until tired
		3	Work in short spells
		4	Work until bored
		5	Something else- please specify

32.	How do you best learn about a new subject of interest to you?	1 2 3 4 5	Reading about it Attending an evening class Joining a club Talking to experts Other - please specify

33.	When working on a short repetitive task do you prefer to work-	1 2 3	Quickly, accepting some errors Steadily, ensuring few errors Slowly, ensuring no errors
34.	When working in groups, it is more important to me that the ask gets done than that the team are happy.	1 2	No Yes
35.	When working in a team, the role I prefer is that of-	1 2 3 4 5	The driving force The 'ideas' person The co-ordinator The workhorse The person who makes sure deadlines are met
36.	Do you have any formal management qualifications	1 2	No Yes - please specify

37.	Are you a member of a learned professional society or institute?	1 2	No Yes - please specify

38.	Have you:-	1 2 3	Ever set up your own business? Considered setting up your own business? Never considered setting up your own business?

This section is collecting information on your interests, activities and hobbies.

39.	What type of music do you prefer?	1 2 3 4 5 6	Do not like music Classical Country Jazz Pop Other - please specify

40.	How do you normally communicate with friends or relatives who do not live locally?	1 2 3	Telephone Letter Other means- please specify
41.	Do you regularly compete in a sport ? (ie more than once per month)	1 2	No Yes
42.	When spending an evening relaxing, I regularly listen to music	1 2	No Yes
43.	What is the main source of your current affairs knowledge?	1 2 3 4 5 6	TV Radio Newspapers Magazines Other people None of these
44.	Indicate any of the following that you have produced/had published or participated in the last 3 years. ***(Mark all that apply)***	1 2 3 4 5 6 7 8	A book An article A letter A computer program A piece of music A photograph A painting or drawing A design
45.	Do you attempt to repair broken mechanical or electrical devices?	1 2 3	Often Occasionally Never
46.	Have you been involved with sports or social club management within the past 3 years?	1 2	No Yes

The final section is collecting information about how you perceive yourself in relation to colleagues who are doing a comparable job or following a similar course of study.

47.	Which of the following statements best describe you?	1 2 3	More a talker than a listener As much a talker as a listener More a listener than a talker
48.	In terms of your general standards, would you describe yourself as a:-	1 2 3 4	Very tidy person Tidy person Reasonably tidy person Rather untidy person

49.	Would you say you are:-	1	More a doer than a 'thinker'
		2	As much a 'doer' as a 'thinker
		3	More a 'thinker' than a 'doer'
50.	How would you describe yourself?	1	Very patient
		2	Patient
		3	Slightly impatient
		4	Impatient
51.	Compared with other people, I would describe myself as-	1	A very imaginative thinker
		2	An imaginative thinker
		3	A practical thinker
		4	A very practical thinker
52.	Compared with other people, would you say you are a-	1	Very private person
		2	Private person
		3	Sociable person
		4	Very sociable person
53.	Compared with other people, how do you rate your problem solving ability?	1	Very good
		2	Good
		3	Quite good
		4	A little weak
54.	Compared with other people how do you rate your artistic ability?	1	Very good
		2	Good
		3	Quite good
		4	A little weak
55.	Compared with other people I know, I would say that I am-	1	Cautious
		2	Quite cautious
		3	Carefree
		4	Very carefree
56.	Compared with other people, how do you rate your mathematical ability?	1	Very good
		2	Good
		3	Quite good
		4	A little weak
57.	When involved in a disagreement I tend to-	1	Keep arguing until I get my own way
		2	Keep arguing until a compromise is reached
		3	Eventually give up
		4	Keep quiet

58.	I tend to be happiest when I am-	1	Very busy
		2	Quite busy
		3	Have some time to sit and think
		4	Have lots of time to sit and think
59.	Do you worry about mistakes you have made?	1	Very often
		2	Often
		3	Occasionally
60.	Most of the time I am bursting with energy.	1	No
		2	Yes
61.	Please indicate the type of work you find most effective for your learning	1	Loosely supervised project work
		2	Lectures and essays on set topics
		3	Personal tuition
		4	Group work and discussion
62.	How important is job security to you?	1	Very important
		2	Important
		3	Unimportant
63.	Given the opportunity to do so would you work-	1	Full-time
		2	Part-time
		3	Casual hours
		4	Not at all
64.	Are you-	1	Very ambitious
		2	Ambitious
		3	Comfortable as you are
65.	How important is it that you get recognition for success in your field?	1	Very important
		2	Important
		3	Not so important
		4	Not important at all
66.	Do you consider yourself to be-	1	More of a risk taker than most
		2	As much of a risk taker as most
		3	Less of a risk taker than most
		4	Not a risk taker

THANK YOU FOR TAKING THE TIME TO COMPLETE THE QUESTIONNAIRE

Enquiries to: Melanie Henderson
University of Northumbria at Newcastle
NE1 8ST
Email melanie.henderson@unn.ac.uk
Tel: 0191 227 3038
Fax: 0191 227 4684

Appendix III

Research brief circulated to organizations

Research Brief: An investigation into the application of biodata in managerial selection. (In association with the University of Northumbria at Newcastle)

Background

Melanie Henderson is currently engaged in a doctoral research programme with the University of Northumbria at Newcastle. The research project is focused upon the use of biographical information (Biodata) in managerial selection. The focus of the research is to test out the ability of certain biographical information to predict standards of managers' work performance and a number of the managers' personal attributes. These include aspects of personality, in particular extroversion-introversion and stability-neuroticism, and critical thinking ability. To date, large samples of part-time mature students on managerial courses, and full time business studies undergraduates have been used to confirm that the biodata questionnaire (adapted from an original design by Wilkinson, 1995), can predict consistently these respondent's scores on psychometric tests of personality (Eysenck Personality Inventory) and a Critical Thinking Ability test (Watson Glaser Critical Thinking Appraisal). So far the results have been highly encouraging.

Dr Larry Wilkinson and Nigel van Zwanenberg of the University of Northumbria are acting in a supervisory capacity to Melanie Henderson

Next steps

After conducting the initial study (involving students of the University of Northumbria) and the literature review, the spring and summer period (until October/November 1998) will be devoted to researching the issues of the nature of managerial work and managerial performance appraisal. Melanie Henderson intends to test out the predictive power of the biodata measures for a total preferably of two to three hundred managers. This will involve gathering performance data, probably using the measures adopted by organisation(s) and possibly a more generic measure as a result of Melanie's research; biographical data, obtained via a questionnaire and data on personality and critical thinking, gathered by the psychometric tests described above.
The research is concerned with trying to predict how well individuals perform in their jobs as managers. The evidence of the predictive power of biodata in this respect is very strong.

It is proposed that the research format could take shape as follows:
Each participant would be required to complete a biodata inventory (10 minutes), the Eysenck Personality Inventory (15 minutes) and the Watson Glaser Critical Thinking Appraisal test (45 minutes). In total, access to each participant for 75 minutes would be needed to complete the test inventories and discuss the uses of biodata in managerial selection.

In addition, access to performance data is required in order to calculate how well the biodata and psychometric test measures can predict managerial success. All candidates will be given a code name and will not be known by name to this researcher or to the research supervisors. Complete confidentiality of personnel records and test information is assured.

How _____could help

The research is dependent on gaining the required data, preferably from a relatively small number of organisations. I believe that _____ would be an excellent source of such data, not least because of the nature of the work undertaken by managerial staff there, and the number of people involved.

What _____ might gain

The information produced from the research of Melanie Henderson may be of assistance in performance measures reviews or investigations. The PhD programme supervisors have expertise in these and related areas and are willing to make this available as a quid pro quo. If, as hoped, the biodata instrument is shown to be a consistent and valid predictor of performance, there are potentially considerable benefits to be gained in the fields of selection into management at _____(and other organisations), in management development and placement.

Additionally, there are likely to be spin-offs for all the parties from taking part in a structured exercise.

What UNN may gain

Apart from the obvious benefit of access to a large group of potential respondents, we get to test out ideas in practice and to add to our appreciation of the day to day issues confronting managers.

How to move the project on

Melanie Henderson would be more than happy to talk to the appropriate people at _____ _____.

Contact:

E-mail
Melanie.Henderson@unn.ac.uk
Tel 0191 227 3038
Fax 0191 227 3682

Appendix IV
Managerial biodata questionnaire

Code

Please make one response per question unless otherwise indicated.

This first section is collecting general background information.

1. Please indicate your sex
   1 Female
   2 Male

2. Indicate your age group
   1 Under 18
   2 18 - 19
   3 20 - 24
   4 25 - 29
   5 30 - 39
   6 40 - 49
   7 over 50

3. What was your family status in your parental home?
   1 Only child
   2 Oldest of several
   3 Youngest of several
   4 Middle child

4. Indicate the nature of the main wage earner's occupation in your parental home.
   1 Professional
   2 Skilled
   3 Semi skilled
   4 Unskilled

5. Please indicate by attributing a % amount, the importance of each of the following in relation to your job -

Category	Importance to the job (%)
Work with People	
Work with Data/ Information	
Work with objects / machinery	
Total	100%

6. Are you left handed or right handed?
   1 Left handed
   2 Right handed

The next section is collecting information about your educational background.

7. When at school, which subject were you most interested in?
   1 Physics
   2 Chemistry or biology
   3 Technology or design
   4 Computing or IT
   5 Mathematics
   6 History
   7 Languages

Are you continuing your formal education? ***(If 'Yes', please go on to Question 8a).***

1 No
2 Yes

8a. Please indicate your main reasons for continuing your education ***(mark all that apply)***

1 Career prospects
2 Challenge and responsibility
3 Job security
4 Interesting and varied work
5 To make best use of my skills
6 Like creating things
7 Like solving practical problems

9. How many 'GCSE' passes, at grade C or better, or equivalents (Scottish 'O' grades) do you have?

1 Less than 3
2 3 - 6
3 7 - 10
4 More than 10

Please answer just one of questions 10a or 10b

10a. How many 'A' level passes do you have?

1 0
2 1
3 2
4 3
5 4 or more

10b. How many Scottish Highers do you have?

1 0
2 1 - 2
3 3 – 4
4 5 - 6
5 7 or more

10c. Do you hold-

1 GNVQ
2 National Certificate or Diploma
3 Higher National Cert. or Diploma
4 Other e.g. International Baccalaureate

11. At what type of institution did you study for your University entry qualifications (e.g. 'A' levels etc.)?

1 Not applicable
2 Secondary School
3 VI Form College
4 Technical or FE College
5 Polytechnic or University

12. Indicate where your pre-university education took place.

1 Scotland
2 England, Wales, Northern Ireland
3 European Community (excl. UK)
4 Europe (excl. European Comm.)
5 Outside Europe

13.	How do you normally revise for your examinations?	1	Plan to cover everything
		2	Concentrate on the weak areas
		3	Prepare selected topics
		4	Cover as much as you can in the time available.
14.	At which of the following have you completed a full course of study? ***(mark all that apply)***	1	University
		2	Polytechnic
		3	College of HE
		4	None of these
15.	Indicate the nature of your highest formal educational qualification.	1	Doctorate
		2	Masters/PG Dip
		3	Degree
		4	Diploma/Certificate
		5	A levels or equivalents
		6	University foundation course

Have you ever studied for pleasure at part-time or evening classes?

1	No
2	Yes

Have you ever studied statistics?

1	No
2	Yes

This section is concerned with how you work on tasks

18.	Which of the following best describes your work style?	1	Work on one task at a time
		2	Work on several tasks at once
		3	Work on many tasks at once
19.	In a work situation I prefer to-	1	Work independently
		2	Work within a loose framework
		3	Work in a well defined framework

Which of the following best describes your approach to work scheduling?

1	Tackle tasks as they arise
2	Tackle tasks in order of work priority
3	Tackle the easy or short tasks first.

21.	What is the size of the project group that you work best with?	1	Large team
		2	Small team
		3	Group of 2 or 3 people
		4	Individual
22.	In which of the following media do you prefer to work?	1	Words
		2	Pictures
		3	Numbers/symbols
		4	No preference

23. Which of the following activities best describes your approach to an unfamiliar problem?

1 Guess a solution and check it works
2 Eliminate non-feasible solutions
3 Attempt to deduce the solution from the information given
4 Attempt to relate to a similar known problem

When working on a problem do you prefer to-

1 Break it up into small stages, solving each stage in turn
2 Attempt to solve the problem in its entirety

25. When working on a project I prefer the situation to-

1 Be very well specified and predictable
2 Have a degree of predictability
3 Have a degree of uncertainty
4 Be uncertain and unpredictable

26. When working on a problem, I like to -

1 Tackle it in logical stages
2 Focus on specific aspects that interests me
3 Explore a range of different approaches
4 Sit back and think of the best way of approaching it
5 Come to an answer intuitively

When things get really difficult I-

1 Carry on regardless
2 Try to find a way around the problem
3 Move onto something else

28. When acquiring a new work skill do you practice until-

1 Nearly perfect
2 Adequate
3 Bored with practice
4 Something else- please specify

29. When learning to use an unfamiliar piece of complicated equipment do you-

1 Read the instruction manual
2 Experiment with the controls
3 Get someone to demonstrate it
4 Other - please specify

30. Which of the following best describes your needs when undertaking a new activity?

1 Careful supervision
2 Some supervision
3 Minimal supervision
4 Other - please specify

31. What is your strategy when

1 Work at it until complete

working on a long and demanding task?

2 Work until tired
3 Work in short spells
4 Work until bored
5 Something else- please specify

32. How do you best learn about a new subject of interest to you?

1 Reading about it
2 Attending an evening class
3 Joining a club
4 Talking to experts
5 Other - please specify

33. When working on a short repetitive task do you prefer to work-

1 Quickly, accepting some errors
2 Steadily, ensuring few errors
3 Slowly, ensuring no errors

When working in groups, it is more important to me that the task gets done that that the team are happy.

1 No
2 Yes

35. When working in a team, the role I prefer is that of-

1 The driving force
2 The 'ideas' person
3 The co-ordinator
4 The workhorse
5 The person who makes sure deadlines are met

36. Do you have any formal management qualifications

1 No
2 Yes - please specify

Are you a member of a learned professional society or institute?

1 No
2 Yes - please specify

38. Have you:-

1 Ever set up your own business?
2 Considered setting up your own business?
3 Never considered setting up your own business?

This section is collecting information on your interests, activities and hobbies.

39. What type of music do you prefer?

1 Do not like music
2 Classical
3 Country
4 Jazz
5 Pop
6 Other - please specify

40.	How do you normally communicate with friends or relatives who do not live locally?	1 2 3	Telephone Letter Other means- please specify
41.	Do you regularly compete in a sport ? (ie more than once per month)	1 2	No Yes
42.	When spending an evening relaxing, I regularly listen to music	1 2	No Yes
43.	What is the main source of your current affairs knowledge?	1 2 3 4 5 6	TV Radio Newspapers Magazines Other people None of these
44.	Indicate any of the following that you have produced/had published or participated in the last 3 years. ***(Mark all that apply)***	1 2 3 4 5 6 7 8	A book An article A letter A computer program A piece of music A photograph A painting or drawing A design
45.	Do you attempt to repair broken mechanical or electrical devices?	1 2 3	Often Occasionally Never
46.	Have you been involved with sports or social club management within the past 3 years?	1 2	No Yes

The final section is collecting information about how you perceive yourself in relation to colleagues who are doing a comparable job or following a similar course of study.

47.	Which of the following statements best describe you?	1 2 3	More a talker than a listener As much a talker as a listener More a listener than a talker
48.	In terms of your general standards, would you describe yourself as a:-	1 2 3 4	Very tidy person Tidy person Reasonably tidy person Rather untidy person

49.	Would you say you are:-	1	More a doer than a 'thinker'
		2	As much a 'doer' as a 'thinker
		3	More a 'thinker' than a 'doer'
50.	How would you describe yourself?	1	Very patient
		2	Patient
		3	Slightly impatient
		4	Impatient
51.	Compared with other people, I would describe myself as-	1	A very imaginative thinker
		2	An imaginative thinker
		3	A practical thinker
		4	A very practical thinker
52.	Compared with other people, would you say you are a-	1	Very private person
		2	Private person
		3	Sociable person
		4	Very sociable person
53.	Compared with other people, how do you rate your problem solving ability?	1	Very good
		2	Good
		3	Quite good
		4	A little weak
54.	Compared with other people how do you rate your artistic ability?	1	Very good
		2	Good
		3	Quite good
		4	A little weak
55.	Compared with other people I know, I would say that I am-	1	Cautious
		2	Quite cautious
		3	Carefree
		4	Very carefree
56.	Compared with other people, how do you rate your mathematical ability?	1	Very good
		2	Good
		3	Quite good
		4	A little weak
57.	When involved in a disagreement I tend to-	1	Keep arguing until I get my own way
		2	Keep arguing until a compromise is reached
		3	Eventually give up
		4	Keep quiet
58.	I tend to be happiest	1	Very busy

	when I am-	2	Quite busy
		3	Have some time to sit and think
		4	Have lots of time to sit and think
59.	Do you worry about mistakes you have made?	1	Very often
		2	Often
		3	Occasionally
60.	Most of the time I am bursting with energy.	1	No
		2	Yes
61.	Please indicate the type of work you find most effective for your learning	1	Loosely supervised project work
		2	Lectures and essays on set topics
		3	Personal tuition
		4	Group work and discussion
62.	How important is job security to you?	1	Very important
		2	Important
		3	Unimportant
63.	Given the opportunity to do so would you work-	1	Full-time
		2	Part-time
		3	Casual hours
		4	Not at all
64.	Are you-	1	Very ambitious
		2	Ambitious
		3	Comfortable as you are
65.	How important is it that you get recognition for success in your field?	1	Very important
		2	Important
		3	Not so important
		4	Not important at all
66.	Do you consider yourself to be-	1	More of a risk taker than most
		2	As much of a risk taker as most
		3	Less of a risk taker than most
		4	Not a risk taker

THANK YOU FOR TAKING THE TIME TO COMPLETE THE QUESTIONNAIRE

Enquiries to:
Melanie Henderson
Tel: 0191 227 3038
Fax: 0191 227 4684
University of Northumbria
NE1 8ST
Email melanie.henderson@unn.ac.uk

Appendix V
Psychometric test feedback report

Code Number _____

Extroversion **Neuroticism** **Lie scale**

WHAT DOES YOUR SCORE MEAN?

An **extrovert** will score above 12 out of a possible 24 items on the Extroversion scale. Many people will score between 10 and 14 on this scale showing mixed tendencies towards both extroversion and introversion ('ambiverts'). An **introvert** will score below 10 on the extroversion scale.

A score above 12 out of a possible score of 24 on the Neuroticism scale indicates that the respondent is somewhat **neurotic** or unstable in their emotions. Individuals who a are more **stable** in their emotional responses tend to score below 12.

The **lie scale** is a measure of **socially desirable responding**. Certain items, such as, 'have you ever been late for an appointment or work' assume that anyone answering positively, is responding in a socially desirable way i.e. 'faking good' consciously or unconsciously for the purposes of the test. The maximum score on this scale is 9.

BACKGROUND TO THE EPI.

The EPI measures the personality constructs of Extroversion/Introversion and Neuroticism/ Stability. Each of these major personality factors can best be illustrated by reference to a continuum:

Introversion.....**Extroversion**

Stability..... **Neuroticism**

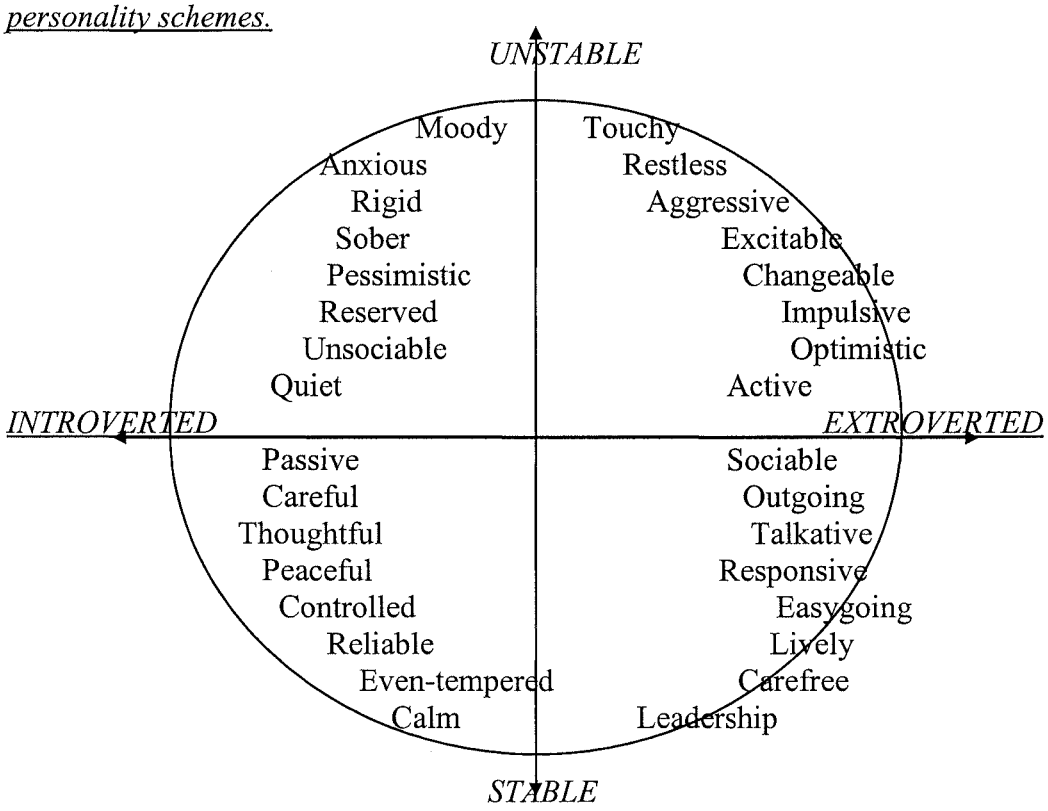
Thus, the more extroverted you are, the less introverted you will be; the more neurotic you are, the less stable your emotional responses will be.

The Extrovert	**The Introvert**
sociable with many friends	quiet and reserved
impulsive	cautious
sensation seeking	reading and study
risk taker	has few but close friends
enjoys parties	plans ahead
carefree	reliable
aggressive	has high ethical standards
feelings not under tight control	feelings under control
unreliable	deals easily with concepts
needs to have experience of things first hand	and ideas

The Neurotic **Emotional stability**

Interpretation of the Neuroticism/stability score is somewhat more problematic than the Extroversion/Introversion scale. It is harder to provide a description of the characteristics of the neurotic responder because the profile will depend upon the scores in terms of Extroversion and Introversion.

Figure 1 Relationship of Extroversion/Introversion and Neuroticism/Stability to other personality schemes.



The Watson-Glaser Critical Thinking Appraisal

The Watson Glaser Critical Thinking Appraisal is a measure of the ability to reason analytically. This test is often used in managerial selection. Five subtests measure the following abilities, considered to represent inter-related but different aspects of critical thinking ability:

Items ask candidates to reason critically about statements involving 'neutral' topics such as the weather, scientific facts or experiments and other matters about which people usually do not have strong feelings or prejudices. Other items, approximately parallel in logical structure, pertain to political, economic and social issues about which people may have definite emotional feelings, biases or prejudices. Inclusion of controversial material is intended to provide a partial sample of an individual's ability to deal critically with issues which may be surrounded by strong feelings or biases. Strong attitudes, biases and opinions have been shown to affect the ability of some
people to think critically (e.g. Jaeger and Freijo, 1975; Jones and Cook, 1975; Mitchell and Byrne, 1973; Sherif, Sherif, and Nebergall, 1965).

Total score _____

Inference	Evaluating the validity of inferences drawn from factual statements.
Recognition of Assumptions	Identifying unstated assumptions or presuppositions.
Deduction	Determining if certain conclusions follow on from information.
Interpretation	Weighing evidence and deciding if certain conclusions are warranted.
Evaluation of Arguments	Distinguishing strong (and relevant) from weak (irrelevant) arguments.

The total possible score of the test is 80. Each subtest contributes a possible 16 marks to the overall score.

Normative Data:

Sample	Age	Mean Score	Sample size
Senior managers in a public sector organisation	30 - 54	60.43	95
Senior executives and Directors	32 - 65	63.23	64

It should be noted that these norm groups do not contain representatives of ethnic minority groups.

Watson, G., & Glaser, E.M. (1991).*Watson-Glaser Critical Thinking Appraisal British Manual.* Kent: PCL.

Appendix VI
Performance proforma and personnel data forms

Personnel Data Records

The following performance appraisal is intended to verify how well the biodata instrument predicts managerial job performance. Thank you for your continued participation in this research.
You and the candidate you are rating will remain anonymous in this phase of the research and this information will not be available to anyone other than the researcher.

Information about the candidate

Code: [ ]

Job title / Nature of his/her work _____

Personnel Data:

How long has the candidate worked for your organisation _____

How many promotions has the candidate had in this time? _____

Approximate salary of the candidate _____

Attendance record:

Absence (excluding annual leave) in the last 6 months exceeds:

no absence	1
1 day - 1 week	2
1 - 3 weeks	3
1 month - 2 months	4
More than 2 months	5

Performance Appraisal

Candidate Code: ☐

Overall Performance Rating

Please rate the overall job performance of the candidate by ticking the appropriate box. (1 represents the lowest level of performance, 6 represents the highest level of performance).

1☐ 2☐ 3☐ 4☐ 5☐ 6☐

Please show the importance of each of the following in relation to the candidate's job

This should be done by allocating a percentage to each category of work so that the total amounts to 100%. Then demonstrate how well you consider the candidate fulfils each of these categories of work by allocating a percentage to reflect the performance of that candidate.

Category	Importance to the job (%)	How well does the candidate meet this requirement of the job?
Work with people		out of 100%
Work with Data / Information		out of 100%
Work with objects (e.g. machinery / equipment)		out of 100%
Total	100%	

Ratings of Performance

In the following section please indicate how you rate the candidate's performance on each dimension by circling a number between 1 and 4 (1 represents the lowest level of performance, 4 represents the highest level of performance). ***Please rate the following***

Work with people

	Low			*High*
How well does the candidate represent the organisation to the public and / or customers	1☐	2☐	3☐	4☐
Rate the quality of the supervision the candidate provides for subordinates	1☐	2☐	3☐	4☐
Can this person organise effective teams that get the job done	1☐	2☐	3☐	4☐
Can this candidate communicate ideas effectively	1☐	2☐	3☐	4☐

Work with data	*Low*			*High*
Candidate's work-related knowledge	1	2	3	4
Candidate's problem solving ability	1	2	3	4
Planning and organisational ability	1	2	3	4

Work with things	*Low*			*High*
Is this candidate proficient in use of equipment required for the job	1	2	3	4
Does this candidate demonstrate knowledge of equipment used by subordinates / peers	1	2	3	4
How well does this candidate allocate resources	1	2	3	4
Does the candidate keep abreast of technical innovations at the workplace	1	2	3	4

General motivation and ability	*Low*			*High*
How dedicated to the organisation is this candidate	1	2	3	4
Is this candidate a conscientious worker	1	2	3	4
How much enthusiasm does this person show consistently?	1	2	3	4
How much energy does this person consistently show?	1	2	3	4
Do you consider this candidate to have promotion potential	1	2	3	4

***Appraisers name / signature*_____**

Please note this section is entirely voluntary.

If you are rating more than one participant of the research, it is not necessary to complete this part of the form more than once, your signature will identify you to the researcher.

For the purposes of further research, it is interesting to collect information on the rater of each participant in order to analyse potential variance attributable to rater - ratee characteristics, if you are willing, this information would be very useful.

Information about you

Your organisational relationship to the candidate _____

Gender	M	1
	F	2
Age	18 - 20	1
	20 - 24	2
	25 - 29	3
	30 - 39	4
	40 - 49	5
	over 50	6
Ethnic identity:	White	1
	Black Caribbean	2
	Black African	3
	Black-other	4
	Indian	5
	Pakistani	6
	Bangladeshi	7
	Chinese	8
	Other	9

Thank you for your time and co-operation.

Appendix VII

Biodata analogue models (Managerial job incumbents)

Significant correlations with Critical Thinking Ability

Item		Correlation	n	P =
Q5a	Work with people	-.182	134	.035
Q5b	Work with data/information	.176	134	.042
Q14-4	Have not completed an HE qualification	-.255	135	.003
Q15	Increasing qualification (degree – doctorate)	-.367	120	.000
Q17	Have studied statistics	.203	134	.018
Q53	Rate problem solving as a little weak	-.208	135	.016
Q28	When acquiring a new skill you practice until perfect	.170	135	.049
Q35	When working in a team, the role I prefer is making sure deadlines are met / workhorse.	-.189	133	.029
Q37	Are you a member of a learned professional society or institute?	.196	135	.023
Q48	By your general standards would you describe yourself as a very tidy person?	.172	135	.046
Q52	Compared with others are you a private person?	-.179	135	.038
Q30-2	Need some supervision when undertaking a new task.	-.182	135	.035
q30-3	Need minimal supervision when undertaking a new activity	.189	135	.028
Q56	How do you rate your mathematical ability (good)	-.330	135	.000
Q62	How important is job security to you (unimportant – very important)	-.241	135	.005
Q40-1	Use telephone to communicate with friends / relatives who do not live locally.	.184	134	.033
Q29-2	Experiment with the controls when learning to use an unfamiliar piece of equipment.	-.208	135	.016
Q7a	At school you were most interested in physics.	.170	133	.050
Q32-1	Reading is the best way for you to learn about a new subject.	.176	135	.041

Critical Thinking weighted item model

COMPUTE WG2wi = -.182 * q5A + .176 * q5B - .255 * q14_4 - .367 * q15 + .203 * q17 + .170 * q28 - .189 * q35 + .196 * q37 + .172 * q48 - .179 * q52 - .208 * q53 - .330 * q56 - .241 * q62 + .184 * q40_1 - .208 * q29_2 + .196 * Larry37 + .170 * q7a - .182 * q30_2 + .189 * q30_3 + .176 * q32_1. EXECUTE .

Rational / unit-weighted model

COMPUTE WG2rat = -1 *q5A + 1 * q5B - 1* q14_4 -1 * q15 + 1* q17 + 1 * q28 - 1 * q35 + 1 * q37+ 1 * q48 - 1 * q52 - 1 * q53 - 1 * q56 - 1 * q62 + 1 * q40_1 - 1 * q29_2 + 1 * Larry37 + 1* q7a - 1 * q30_2 + 1 * q30_3 + 1 * q32_1.EXECUTE .

Item	Items rejected from models due to low response rate (less than 10% of all responses to the item)	r	n	P =	Response rate
Q44-1	I have published a book in the last 3 years	.171	135	.047	5.9% said yes
Q40-2	Write letters telephone to communicate with friends / relatives who do not live locally.	-.191	134	.027	6.7% said yes
Q11-5	You studied a foundation course at Polytechnic for entry into HE.	-.194	131	.027	8.1% said yes

Appendix VI
Biodata analogue model of extroversion (Managerial job incumbents)

Significant correlations with Extroversion

Item		Correlation	n	P =
Q38	I have never considered setting up my own business	.177	135	.040
Q41	Regularly compete in a sport	.185	135	.032
Q42	Regularly listen to music	.183	135	.034
Q47	More a listener than a talker	-.343	135	.000
Q49	More a thinker than a doer	-.258	135	.003
Q52	Compared with others, I would describe myself as a very sociable person	.525	135	.000
Q50	How would you describe yourself? Impatient – patient	.208	135	.016
Q51	I would describe myself as a very practical thinker	-.276	135	.001
Q55	Compared with other people, I would say that I am very carefree	.419	135	.000
Q57	When involved in a disagreement, I tend to keep quiet	-.254	134	.003
Q58	I am happiest with lots of time to sit and think – very busy	-.170	135	.049
Q59	Do you worry about mistakes you have made? Very occasionally – often.	.236	135	.006
Q60	Most of the time I am bursting with energy	.418	135	.000
Q64	I am ambitious - comfortable as you are (very ambitious)	-.307	135	.000
Q66	I consider myself to be ...more of a risk taker than most - not a risk taker	-.458	135	.000
Q62	Job security is unimportant to me	.177	135	.000
Q61-4	Group work and discussion are most effective for my learning	.178	134	.040
Q61-2	I find lectures and essays on set topics most effective for my learning	-.203	134	.019
Q7	You are a member of a professional of learned society or institute	-.189	135	.029
Q36	You have formal management qualifications	.171	134	.048
Q39-2	Prefer classical music	-.142	135	.001
Q21	I work best with a project group that is a team (Large team – individual)	-.172	134	.013
Q3	Status in the parental home (only child / oldest of several)	.202	132	.020
Q2	Increasing age	-.231	135	.007
Q14-4	Have not completed an HE qualification	.203	135	.018
Q20	Prefer to tackle easy or short tasks first	-.195	135	.024
Q25	I prefer projects to have a degree of specificity and predictability (versus uncertainty)	.214	135	.013

Extroversion weighted item model

COMPUTE Ex2wi = -.189 * larry37 + .171 * larry36 - .231 * larry2 - .142 * q39_2 - .203 * q61_2 + .178 * q61_4+ .419 * q55 - .254 * q57-.170 * q58 + .236 * q59 + .418 * q60 + .177 * q62 - .307 * q64 - .458 * q66 + .202 * q3 - .170 * q7 + .203 * q14_4 - .195 * q20 - .172 * q21 + .214 * q25 + .177 * q38 + .272 * q39 + .185* q41 + .183 * q42 - .343 * q47 - .258 * q49 + .208 * q50 - .276 * q51 + .525 * q52.EXECUTE .

Extroversion rational / unit weighted model

COMPUTE Ex2rat = -1 * larry37 + 1 * larry36 - 1 * larry2 - 1 * q39_2 - 1 * q61_2 + 1* q61_4+ 1 * q55 - 1 * q57 - 1 * q58 + 1 * q59 + 1 * q60 + 1 * q62 - 1 * q64 - 1 * q66 + 1 * q3 - 1 * q7 + 1 * q14_4 - 1 * q20 - 1 * q21 + 1 * q25 + 1 * q38 + 1* q39 + 1* q41 + 1 * q42 - 1 * q47 - 1 * q49 + 1 * q50 - 1 * q51 + 1 * q52.EXECUTE .

Item	Items rejected from models due to low response rate (less than 10% of all responses to the item)	r	n	P =	Response rate
Q10c-1	Hold a GNVQ certificate	.202	135	.019	3.7% said yes
Q7-2	Preferred chemistry or biology at school	.192	133	.027	8.9% said yes

Appendix VI

Biodata analogue model of neuroticism (Managerial job incumbents)

Significant correlations with Neuroticism

Item		Correlation	n	P =
Q59	Do you worry about mistakes you have made very often - occasionally (often)	-.403	135	.000
Q8	Are you currently studying for a HE qualification	-.231	135	.007
Q25	Preference for project work (well specified - uncertain and unpredictable) prefer more predictable project work	-.219	135	.011
Q62	How important is job security to you? Very important - unimportant (very important)	-.269	135	.002
Q55	Compared with others I am cautious - carefree (cautious)	-.281	135	.001
Q18	Work style on one task at a time -many tasks at once (one at a time)	-.231	135	.007
Q8a-2	The main reason I am continuing my education is for challenge and responsibility	-.171	135	.069
Q8a-5	The main reason I am continuing my education is to make best use of my skills	-.241	135	.005
Q10c-2	I have a National certificate or diploma	-.170	135	.048
Q14-2	I have completed a full course of study at a polytechnic	.170	135	.048
Q23_2	Prefer to eliminate non-feasible solutions when working on an unfamiliar problem	-.186	135	.031
Q44-2	Have published an article in the last 3 months	-.203	135	.018
Q50	How would you describe yourself Patient - impatient (Impatient)	.171	135	.048
Q23-4	When approaching an unfamiliar problem attempt to relate to a similar known problem	.194	135	.024
Q60	Most of the time I am bursting with energy (no)	-.326	135	.000
Q52	Compared to others I am a private person	-.239	135	.005
Q57	When involved in an argument I tend to give up / keep quiet	.221	134	.010
Q64	In terms of ambition – comfortable as you are	.384	135	.000
Q63	Given the choice I would prefer to work casually / not at all	.198	.135	.021
Q29-3	When learning how to use unfamiliar piece of equipment I prefer to get someone to demonstrate it	.172	135	.047
Q7-4	Preference for computing or IT while at school	-.204	133	.019
Q35-4	In teams I prefer the role of 'workhorse'	.219	133	.011
Q39-2	Preference for classical music	.188	135	.029
Q61-3	I find personal tuition most effective for my learning	.209	134	.015

Weighted item model

COMPUTE N2wi = -.231 * q8 - .171 * q8a_2 - .241 * q8_5 - .170 * q10c_2 + .170 * q14_2 - .231 * q18 - .219 * q25 - .203 * q44_2 + .171 * q50 - .239 * q52 - .281 * q55 + .221 * q57 - .403 * q59 - .326 * q60 - .269 * q62 +.384 * q64+ .198 * q63 - .171 * q8_2 + .172 * q29_3 - .186 * q23_3 + .194 * q23_4 + .219 * q35_4 + .188 * q39_2.EXECUTE .

Rational / unit weighted model

COMPUTE N2rat = -1 * q8 - 1 * q8a_2 - 1 * q8_5 - 1 * q10c_2 + 1 * q14_2 - 1 * q18 - 1 * q25 - 1 * q44_2 + 1 * q50 - 1 * q52 - 1 * q55 + 1 * q57 - 1 * q59 - 1 * q60 - 1 * q62 + 1 * q64+ 1 * q63 - 1 * q8_2 + 1 * q29_3 - 1 * q23_3 + 1 * q23_4 + 1 * q35_4 + 1 * q39_2. EXECUTE .

Appendix VII

Common items among biodata analogue models

Item		Critical thinking	Neuroticism	Extroversion	Energy*	Conscientiousness*
Gender					**female**	**female**
Age				-	-	-
Q3	Family status			+	+	
Q5a	**Work with people**	+			+	
Q5b	**Work with data**	+			-	
Q7a	Prefer physics	+				
Q7-5	Prefer maths					-
Q8	Continuing education		-			
Q8-2	Challenge & responsibility		-			
Q8-5	To make best use of skills		-		+	
Q10c-2	National certificate/dip.		-			
Q10c-3	HND/HNC					+
Q13-2	Concentrate on weak areas in revision					+
Q14-1	University education					-
Q14-2	Polytechnic education		+			
Q14-4	**No higher education qualifications**	-		+		
Q15	Highest qualification	-				
Q16	**Evening classes**				+	+
Q17	Studied statistics	+				
Q18	Work on one-many tasks		-			
Q20	Work scheduling			-		
Q21	Size of project group			-		
Q21-2	Work with small team				+	
Q23	Unfamiliar problem					+
Q23-2	Eliminate non-feasible solutions		-			
Q23-4	**Relate to known problem**		+			+
Q24	**Solve problem in entirety**				-	-
Q25	**Predictable-unpredictable work**		-	+		
Q26-1	Tackle in logical stages				+	
Q28	Practice new skills	+				
Q28-2	Practice until adequate				-	
Q29-2	Experiment with controls	-				
Q29-3	Ask for a demonstration		+			
Q30-2	**Need some supervision**	-				-
Q30-3	Need minimal supervision	+				+
Q32-1	Read about new subjects	+				
Q32-4	Talk to experts				-	
Q35	Driving force – workhorse	-				
Q35-4	Workhorse		+			
Q36	Formal management qualifications			+		
Q37	**Professional society membership**	+		-		
Q38	Set up own business			+		
Q39_2	**Prefer classical music**		+	+		
Q40-1	Communicate by letter	+				
Q41	Sports			+		
Q42	Listen to music			+		
Q46	Social club management					-
Q47	Talker-listener			-		
Q48	Tidy-untidy	+				
Q49	Doer-thinker			-		
Q50	**Patient-impatient**		+	+		
Q51	Imaginative-practical			-		
Q52	**Private-sociable**	-	-	+	+	
Q53	Problem solving good-weak	-				
Q55	**Artistic ability good-weak**		-	+		
Q56	Maths ability good-weak	-				
Q57	**Argue – stay quiet**		-	-		
Q58	Busy-time to think			-		
Q59	**Worry about mistakes**		-	+	+	
Q60	**Bursting with energy**		-	+	+	
Q61-1	Prefer project work				+	
Q61_2	Prefer lectures & essays			-		
Q61-4	**Prefer group work**			+	+	
Q62	**Job security (imp – unimp.)**	-	-	+		
Q63	**Work full-time – not at all**		+		-	
Q64	**Ambitious – comfortable**		+	-	-	
Q66	Risk taker (more-less)			-		
Total		19	21	27	19	13

* denotes supervisor rating of trait rather than psychometric self-report measure

Raw Data

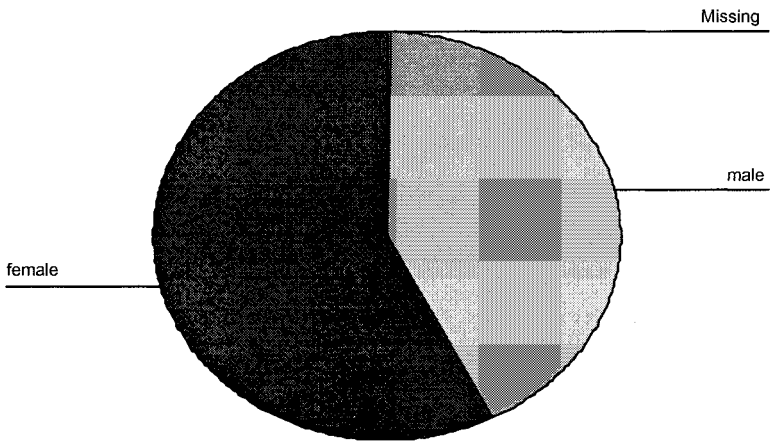
Section 1 Pilot sample statistics

Descriptive statistics

	N	Minimum	Maximum	Mean	Std. Deviation
EXTROV	242	4.00	23.00	13.9256	4.0385
NEUROT	242	.00	24.00	10.5041	4.7425
LIE	242	.00	3.00	1.9711	1.0442
WGSCORE	172	33.00	72.00	56.0058	8.1889
AGE	191	19.00	54.00	28.4293	8.4020
Valid N (listwise)	128				

FEM

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	102	42.0	42.1	42.1
	1.00	140	57.6	57.9	100.0
	Total	242	99.6	100.0	
Missing	System	1	.4		
Total		243	100.0		



AGE * gender Crosstabulation

Count

		gender		
		female	male	Total
AGE	19.00	7	1	8
	20.00	13	2	15
	21.00	6	2	8
	22.00	18	8	26
	23.00	2	17	19
	24.00	5	1	6
	25.00	6	3	9
	26.00	7	5	12
	27.00	6	8	14
	28.00	3	3	6
	29.00	3	2	5
	30.00	3	2	5
	31.00	1	1	2
	32.00	1	1	2
	33.00	1	4	5
	34.00	3	1	4
	35.00	1	3	4
	36.00	1	1	2
	37.00	4	2	6
	38.00	1	2	3
	39.00	2	3	5
	40.00	1	1	2
	41.00	2	4	6
	43.00		2	2
	44.00		1	1
	45.00	1		1
	46.00	1	1	2
	47.00		4	4
	49.00		1	1
	50.00		2	2
	52.00		1	1
	53.00		1	1
	54.00		1	1

EXTROV * gender Crosstabulation

Count

		gender		
		female	male	Total
EXTROV	4.00	2		2
	5.00	3		3
	6.00	1	1	2
	7.00	7	5	12
	8.00	3	2	5
	9.00	6	6	12
	10.00	6	8	14
	11.00	11	6	17
	12.00	12	2	14
	13.00	5	10	15
	14.00	20	12	32
	15.00	16	14	30
	16.00	12	9	21
	17.00	12	5	17
	18.00	6	10	16
	19.00	4	3	7
	20.00	6	2	8
	21.00	4	4	8
	22.00	2	2	4
	23.00	1	1	2
Total		130	102	241

NEUROT * gender Crosstabulation

Count

		gender		
		female	male	Total
NEUROT	.00	1		1
	1.00	2	3	5
	2.00	1	4	5
	3.00	3	3	6
	4.00	6	4	10
	5.00	11	4	15
	6.00	7	5	12
	7.00	4	4	8
	8.00	8	10	18
	9.00	16	8	24
	10.00	8	7	15
	11.00	10	7	17
	12.00	13	12	25
	13.00	11	8	19
	14.00	4	9	13
	15.00	11	4	15
	16.00	5	3	8
	17.00	6	4	10
	18.00	3	1	4
	19.00	1	2	3
	20.00	2		2
	21.00	2		2
	22.00	2		2
	23.00	1		1
	24.00	1		1

WGSCORE * gender Crosstabulation

Count

		gender		
		female	male	Total
WGSCORE	33.00		1	1
	36.00	1		1
	37.00	1		1
	38.00	1		1
	40.00		1	1
	41.00	1	1	2
	42.00	3		3
	43.00	2	1	3
	44.00	2		2
	45.00	3	2	5
	46.00	4	1	5
	47.00	2	2	4
	48.00	5	4	9
	49.00		2	2
	50.00	3	1	4
	51.00	5	1	6
	52.00	3	3	6
	53.00	2	1	3
	54.00	2	2	4
	55.00	8	2	10
	56.00	7	4	11
	57.00	4	5	9
	58.00	8	6	14
	59.00	6	2	8
	60.00	5	2	7
	61.00	3	2	5
	62.00	3	3	6
	63.00	1	2	3
	64.00	4	1	5
	65.00	4	3	7
	66.00	3	5	8
	67.00	2	2	4
	68.00		1	1
	69.00	1	2	3
	70.00	2	2	4
	71.00	2		2
	72.00		1	1

Pilot analogue models

Critical thinking developmental analogue scales

COMPUTE MWWG =.243 * q2 - .261 * q5 + .185 * q9 + .280 * q10a + .173 * q25 - .214 * q53 + .157 * q55 - .161 * q58 - .227 * q23_2 + .156 * q30_3 + .185 * q12_2+ .155 * q44_2 - .202 * q44_3 .EXECUTE .

COMPUTE MratWG =.1 * q2 - 1 * q5 + 1 * q9 + 1 * q10a + 1 * q25 - 1 * q53 + 1 * q55 - 1 * q58 - 1 * q23_2 + 1 * q30_3 + 1 * q12_2+ 1 * q44_2 - 1 * q44_3 .EXECUTE .

CORRELATIONS /VARIABLES=MWIWG MRATWG wgscore /PRINT=TWOTAIL SIG /MISSING=PAIRWISE .

Correlations

		MWIWG	MRATWG	WGSCORE
MWIWG	Pearson Correlation	1.000	.989	.428
	Sig. (2-tailed)	.	.000	.000
	N	227	227	159
MRATWG	Pearson Correlation	.989	1.000	.416
	Sig. (2-tailed)	.000	.	.000
	N	227	227	159
WGSCORE	Pearson Correlation	.428	.416	1.000
	Sig. (2-tailed)	.000	.000	.
	N	159	159	172

Reliability Coefficients

N of Cases = 227.0 N of Items = 13 Alpha = .4604

Extroversion developmental analogue scales

COMPUTE Ex2wi = -.189 * larry37 + .171 * larry36 - .231 * larry2 - .142 * q39_2 - .203 * q61_2 + .178 * q61_4 + .419 * q55 - .254 * q57-.170 * q58 + .236 * q59 + .418 * q60 + .177 * q62 - .307 * q64 - .458 * q66 + .202 * q3 + .203 * q14_4 - .195 * q20 - .172 * q21 + .214 * q25 + .177 * q38 + .185* q41 + .183 * q42 - .343 * q47 - .258 * q49 + .208 * q50 - .276 * q51 + .525 * q52.EXECUTE .

COMPUTE Ex2rat = -1 * larry37 + 1 * larry36 - 1 * larry2 - 1 * q39_2 - 1 * q61_2 + 1* q61_4+ 1 * q55 - 1 * q57 - 1 * q58 + 1 * q59 + 1 * q60 + 1 * q62 - 1 * q64 - 1 * q66 + 1 * q3 + 1 * q14_4 - 1 * q20 - 1 * q21 + 1 * q25 + 1 * q38 + 1* q41 + 1 * q42 - 1 * q47 - 1 * q49 + 1 * q50 - 1 * q51 + 1 * q52. EXECUTE .

Correlations

		MWIEX	MRATEX	EXTROV
MWIEX	Pearson Correlation	1.000	.972	.704
	Sig. (2-tailed)	.	.000	.000
	N	223	223	222
MRATEX	Pearson Correlation	.972	1.000	.673
	Sig. (2-tailed)	.000	.	.000
	N	223	223	222
EXTROV	Pearson Correlation	.704	.673	1.000
	Sig. (2-tailed)	.000	.000	.
	N	222	222	242

Alpha reliability

N of Cases = 223.0 **Reliability Coefficients 28 items**
Alpha = .5885 **Standardized item alpha = .6047**

Neuroticism developmental analogue scales

COMPUTE N2wi = -.231 * q8 - .171 * q8a_2 - .241 * q8_5 - .170 * q10c_2 + .170 * q14_2 - .231 * q18 - .219 * q25 - .203 * q44_2 + .171 * q50 - .239 * q52 - .281 * q55 + .221 * q57 - .403 * q59 - .326 * q60 - .269 * q62 +.384 * q64+ .198 * q63 - .171 * q8_2 + .172 * q29_3 - .186 * q23_3 + .194 * q23_4 + .219 * q35_4 + .188 * q39_2.EXECUTE .

COMPUTE N2rat = -1 * q8 - 1 * q8a_2 - 1 * q8_5 - 1 * q10c_2 + 1 * q14_2 - 1 * q18 + 1 * q23 - 1 * q25- 1 * q44_2 + 1 * q50 - 1 * q52 - 1 * q55 + 1 * q57 - 1 * q59 - 1 * q60 - 1 * q62 + 1 * q64+ 1 * q63 - 1 * q8_2 + 1 * q29_3 - 1 * q23_3 + 1 * q23_4 + 1 * q35_4 + 1 * q39_2. EXECUTE .

		MWIN	MRATN	NEUROT
MWIN	r	1.000	.968	.559
	Sig. (2-tailed)	.	.000	.000
	N	227	227	226
MRATN	r	.968	1.000	.521
	Sig. (2-tailed)	.000	.	.000
	N	227	227	226
NEUROT	r	.559	.521	1.000
	Sig. (2-tailed)	.000	.000	.
	N	226	226	242

Alpha reliability

N of Cases = 227.0 **Reliability Coefficients 17 items**
Alpha = .6213 **Standardized item alpha = .6232**

Test retest reliability

Test-retest data for analogues developed from the managerial sample
(using pilot sample participants)

5% of the original sample completed a second biodata form 2 years after first completion. The biodata analogue models developed from the pilot sample of 'potential managers' are shown to be robust over time.

Correlations

		WGA	WGB
WGA	Pearson Correlation	1.000	.991
	Sig. (2-tailed)	.	.001
	N	5	5
WGB	Pearson Correlation	.991	1.000
	Sig. (2-tailed)	.001	.
	N	5	8

Correlations

		EXA	EXB
EXA	Pearson Correlation	1.000	.955
	Sig. (2-tailed)	.	.003
	N	7	6
EXB	Pearson Correlation	.955	1.000
	Sig. (2-tailed)	.003	.
	N	6	8

Correlations

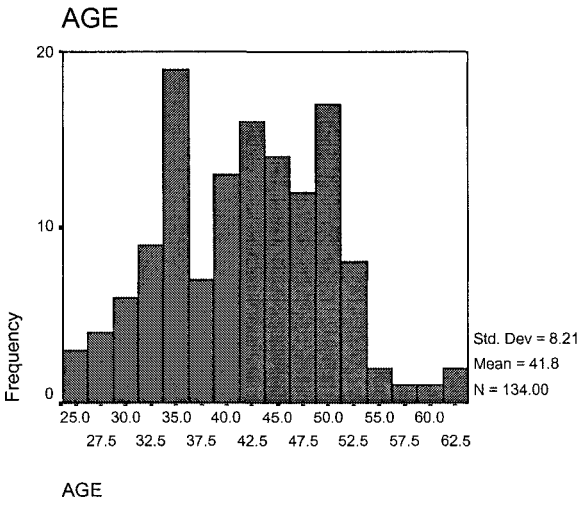
		NA	NB
NA	Pearson Correlation	1.000	.919
	Sig. (2-tailed)	.	.000
	N	9	9
NB	Pearson Correlation	.919	1.000
	Sig. (2-tailed)	.000	.
	N	9	11

Section 2 Managerial sample and incremental validity analysis

	N	Min.	Max.	Mean	Std. Error	Std. Deviation	Variance
EXTROV	135	3.00	23.00	12.2444	.3479	4.0417	16.335
NEUROT	135	.00	21.00	6.7407	.3882	4.5103	20.343
LIE	135	.00	7.00	2.6741	.1497	1.7399	3.027
SEX	135	1.00	2.00	1.1259	2.866E-02	.3330	.111
AGE	134	25.00	63.00	41.7612	.7095	8.2126	67.446
W.G	135	35.00	77.00	59.8074	.7365	8.5575	73.231
Q5A	134	.00	95.00	51.8134	1.5530	17.9769	323.168
Q5B	134	5.00	100.00	35.5152	1.4743	17.0661	291.251
Q5C	134	.00	50.00	12.6714	.9098	10.5311	110.905
TOTALPER	120	2.00	6.00	4.3917	8.052E-02	.8821	.778
IMP1	134	10.00	90.00	46.7013	1.5532	17.9794	323.260
REQ1	122	25.00	100.00	75.0492	1.3617	15.0404	226.212
IMP2	134	5.00	80.00	37.2743	1.3754	15.9219	253.508
REQ2	122	10.00	100.00	75.2295	1.7003	18.7801	352.691
IMP3	134	.00	70.00	16.0243	1.1957	13.8407	191.564
REQ3	107	.00	100.00	74.8505	2.0952	21.6724	469.694
P1	133	2.00	4.00	3.2707	5.661E-02	.6528	.426
P2	120	1.00	4.00	2.8167	6.587E-02	.7216	.521
P3	128	1.00	4.00	2.9297	6.943E-02	.7855	.617
P4	134	1.00	4.00	3.0000	5.297E-02	.6131	.376
D1	134	1.00	4.00	3.2612	5.598E-02	.6481	.420
D2	134	2.00	4.00	3.0299	5.498E-02	.6365	.405
D3	133	1.00	4.00	3.0075	6.269E-02	.7230	.523
T1	131	1.00	4.00	3.2748	6.230E-02	.7131	.509
T2	127	1.00	4.00	3.1417	6.297E-02	.7096	.504
T3	132	1.00	4.00	2.8939	5.717E-02	.6568	.431
T4	130	1.00	4.00	3.0538	6.397E-02	.7294	.532
GM1	134	1.00	4.00	3.3209	6.409E-02	.7419	.550
GM2	134	2.00	4.00	3.5522	5.032E-02	.5825	.339
GM3	134	1.00	4.00	3.2090	6.092E-02	.7052	.497
GM4	134	1.00	4.00	3.1269	6.818E-02	.7892	.623
GM5	133	1.00	4.00	2.9624	7.727E-02	.8911	.794
MAN_SEX	132	1.00	2.00	1.0303	1.498E-02	.1721	2.961E-02
MAN_AGE	129	4.00	6.00	5.3411	6.078E-02	.6903	.477
MAN_ETHN	132	1.00	1.00	1.0000	.0000	.0000	.000
WORKTYPE	0						
WORKTIME	134	.25	27.00	10.8668	.5038	5.8321	34.013
PROMTNS	134	.00	6.00	1.6493	.1255	1.4523	2.109
SALARY	134	10438.00	49300.00	21763.4776	750.4586	8687.1863	75467206.432
ABSENCE	134	1.00	4.00	1.3731	5.962E-02	.6901	.476
CP	134	.00	60.00	14.6328	1.1160	12.9183	166.884
PEOPLE	118	6.00	16.00	12.0593	.2037	2.2122	4.894
DATA	133	6.00	12.00	9.2932	.1244	1.4343	2.057
THINGS	125	5.00	16.00	12.3680	.1926	2.1535	4.638
Valid N (listwise)	0						

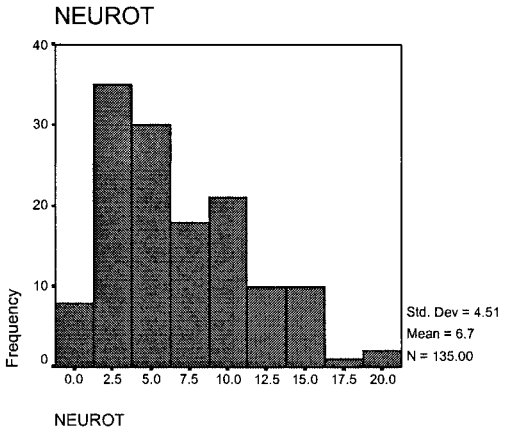
AGE

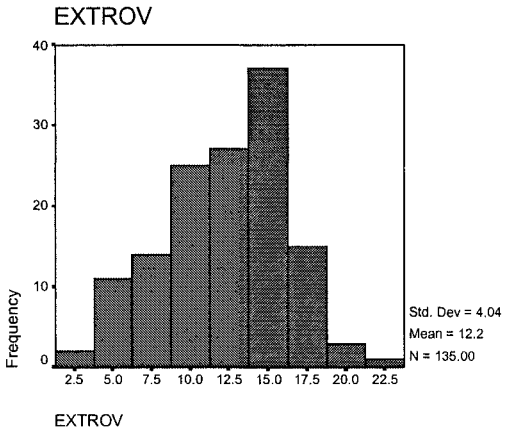
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	25.00	2	1.5	1.5	1.5
	26.00	1	.7	.7	2.2
	27.00	2	1.5	1.5	3.7
	28.00	2	1.5	1.5	5.2
	29.00	4	3.0	3.0	8.2
	30.00	2	1.5	1.5	9.7
	32.00	5	3.7	3.7	13.4
	33.00	4	3.0	3.0	16.4
	34.00	13	9.6	9.7	26.1
	35.00	2	1.5	1.5	27.6
	36.00	4	3.0	3.0	30.6
	37.00	4	3.0	3.0	33.6
	38.00	3	2.2	2.2	35.8
	39.00	5	3.7	3.7	39.6
	40.00	2	1.5	1.5	41.0
	41.00	6	4.4	4.5	45.5
	42.00	7	5.2	5.2	50.7
	43.00	9	6.7	6.7	57.5
	44.00	3	2.2	2.2	59.7
	45.00	7	5.2	5.2	64.9
	46.00	4	3.0	3.0	67.9
	47.00	7	5.2	5.2	73.1
	48.00	5	3.7	3.7	76.9
	49.00	6	4.4	4.5	81.3
	50.00	5	3.7	3.7	85.1
	51.00	6	4.4	4.5	89.6
	52.00	4	3.0	3.0	92.5
	53.00	4	3.0	3.0	95.5
	54.00	2	1.5	1.5	97.0
	58.00	1	.7	.7	97.8
	61.00	1	.7	.7	98.5
	62.00	1	.7	.7	99.3
	63.00	1	.7	.7	100.0
	Total	134	99.3	100.0	
Missing	System	1	.7		



EXTROV

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3.00	2	1.5	1.5	1.5
	4.00	1	.7	.7	2.2
	5.00	3	2.2	2.2	4.4
	6.00	7	5.2	5.2	9.6
	7.00	12	8.9	8.9	18.5
	8.00	2	1.5	1.5	20.0
	9.00	8	5.9	5.9	25.9
	10.00	12	8.9	8.9	34.8
	11.00	5	3.7	3.7	38.5
	12.00	9	6.7	6.7	45.2
	13.00	18	13.3	13.3	58.5
	14.00	13	9.6	9.6	68.1
	15.00	13	9.6	9.6	77.8
	16.00	11	8.1	8.1	85.9
	17.00	8	5.9	5.9	91.9
	18.00	7	5.2	5.2	97.0
	19.00	2	1.5	1.5	98.5
	21.00	1	.7	.7	99.3
	23.00	1	.7	.7	100.0
	Total	135	100.0	100.0	



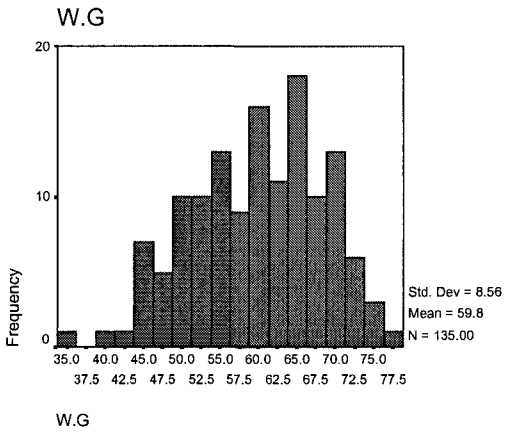


NEUROT

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	3	2.2	2.2	2.2
	1.00	5	3.7	3.7	5.9
	2.00	17	12.6	12.6	18.5
	3.00	18	13.3	13.3	31.9
	4.00	9	6.7	6.7	38.5
	5.00	14	10.4	10.4	48.9
	6.00	7	5.2	5.2	54.1
	7.00	13	9.6	9.6	63.7
	8.00	5	3.7	3.7	67.4
	9.00	7	5.2	5.2	72.6
	10.00	8	5.9	5.9	78.5
	11.00	6	4.4	4.4	83.0
	12.00	6	4.4	4.4	87.4
	13.00	4	3.0	3.0	90.4
	14.00	5	3.7	3.7	94.1
	15.00	2	1.5	1.5	95.6
	16.00	3	2.2	2.2	97.8
	17.00	1	.7	.7	98.5
	19.00	1	.7	.7	99.3
	21.00	1	.7	.7	100.0
	Total	135	100.0	100.0	

W.G

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	35.00	1	.7	.7	.7
	41.00	1	.7	.7	1.5
	42.00	1	.7	.7	2.2
	44.00	2	1.5	1.5	3.7
	45.00	1	.7	.7	4.4
	46.00	4	3.0	3.0	7.4
	47.00	3	2.2	2.2	9.6
	48.00	2	1.5	1.5	11.1
	49.00	1	.7	.7	11.9
	50.00	4	3.0	3.0	14.8
	51.00	5	3.7	3.7	18.5
	52.00	6	4.4	4.4	23.0
	53.00	4	3.0	3.0	25.9
	54.00	5	3.7	3.7	29.6
	55.00	2	1.5	1.5	31.1
	56.00	6	4.4	4.4	35.6
	57.00	5	3.7	3.7	39.3
	58.00	4	3.0	3.0	42.2
	59.00	4	3.0	3.0	45.2
	60.00	6	4.4	4.4	49.6
	61.00	6	4.4	4.4	54.1
	62.00	5	3.7	3.7	57.8
	63.00	6	4.4	4.4	62.2
	64.00	7	5.2	5.2	67.4
	65.00	4	3.0	3.0	70.4
	66.00	7	5.2	5.2	75.6
	67.00	4	3.0	3.0	78.5
	68.00	6	4.4	4.4	83.0
	69.00	6	4.4	4.4	87.4
	70.00	3	2.2	2.2	89.6
	71.00	4	3.0	3.0	92.6
	72.00	2	1.5	1.5	94.1
	73.00	4	3.0	3.0	97.0
	74.00	2	1.5	1.5	98.5
	76.00	1	.7	.7	99.3
	77.00	1	.7	.7	100.0



T-tests of sex and age differences

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	male	118	87.4	87.4	87.4
	female	17	12.6	12.6	100.0
	Total	135	100.0	100.0	

Group Statistics

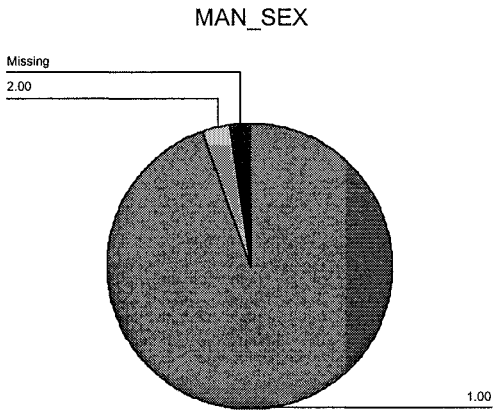
	SEX	N	Mean	Std. Deviation	Std. Error Mean
EXTROV	1.00	118	12.1949	4.1618	.3831
	2.00	17	12.5882	3.1634	.7672
NEUROT	1.00	118	6.7203	4.5176	.4159
	2.00	17	6.8824	4.5946	1.1144
LIE	1.00	118	2.6356	1.7815	.1640
	2.00	17	2.9412	1.4349	.3480
AGE	1.00	117	42.7949	7.5835	.7011
	2.00	17	34.6471	9.0619	2.1978
W.G	1.00	118	60.1102	8.8731	.8168
	2.00	17	57.7059	5.6763	1.3767
TOTALPER	1.00	105	4.3333	.9164	8.943E-02
	2.00	15	4.8000	.4140	.1069
IMP1	1.00	117	47.3047	18.5519	1.7151
	2.00	17	42.5490	13.0719	3.1704
REQ1	1.00	107	73.0935	14.7888	1.4297
	2.00	15	89.0000	7.8376	2.0237
IMP2	1.00	117	37.0635	16.1175	1.4901
	2.00	17	38.7255	14.8783	3.6085
REQ2	1.00	107	74.3738	19.3788	1.8734
	2.00	15	81.3333	12.6246	3.2597
IMP3	1.00	117	15.6319	13.9459	1.2893
	2.00	17	18.7255	13.1700	3.1942
REQ3	1.00	93	74.4516	21.9899	2.2802
	2.00	14	77.5000	19.9759	5.3388
P1	1.00	117	3.2308	.6616	6.116E-02
	2.00	16	3.5625	.5123	.1281
P2	1.00	106	2.7830	.7304	7.094E-02
	2.00	14	3.0714	.6157	.1646
P3	1.00	112	2.9107	.8116	7.668E-02
	2.00	16	3.0625	.5737	.1434
P4	1.00	117	2.9744	.6360	5.880E-02
	2.00	17	3.1765	.3930	9.531E-02
D1	1.00	117	3.2821	.6545	6.051E-02
	2.00	17	3.1176	.6002	.1456
D2	1.00	117	3.0000	.6565	6.070E-02
	2.00	17	3.2353	.4372	.1060
D3	1.00	116	2.9655	.7452	6.919E-02
	2.00	17	3.2941	.4697	.1139
T1	1.00	114	3.2456	.7355	6.888E-02
	2.00	17	3.4706	.5145	.1248
T2	1.00	111	3.0991	.7256	6.887E-02
	2.00	16	3.4375	.5123	.1281
T3	1.00	115	2.8696	.6821	6.361E-02
	2.00	17	3.0588	.4287	.1040
T4	1.00	113	3.0177	.7557	7.109E-02
	2.00	17	3.2941	.4697	.1139
GM1	1.00	117	3.3162	.7503	6.936E-02
	2.00	17	3.3529	.7019	.1702
GM2	1.00	117	3.5128	.5962	5.512E-02
	2.00	17	3.8235	.3930	9.531E-02
GM3	1.00	117	3.1538	.7146	6.606E-02
	2.00	17	3.5882	.5073	.1230
GM4	1.00	117	3.0769	.8004	7.399E-02
	2.00	17	3.4706	.6243	.1514
GM5	1.00	116	2.8793	.9056	8.408E-02
	2.00	17	3.5294	.5145	.1248
CP	1.00	117	14.6887	13.2056	1.2209
	2.00	17	14.2483	11.0784	2.6869
WORKTIME	1.00	117	10.9564	5.9758	.5525
	2.00	17	10.2500	4.8348	1.1726
PROMTNS	1.00	117	1.6581	1.4749	.1364
	2.00	17	1.5882	1.3257	.3215
SALARY	1.00	117	22889.38	8658.0364	800.4358
	2.00	17	14014.65	3283.7127	796.4173

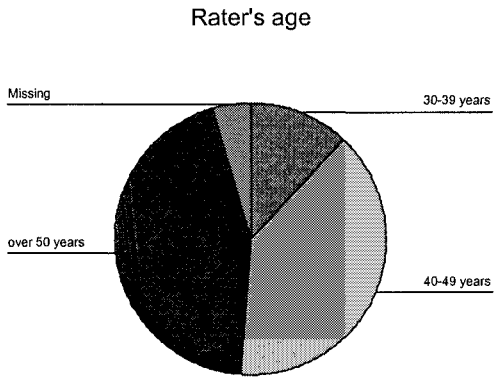
Independent Samples Test

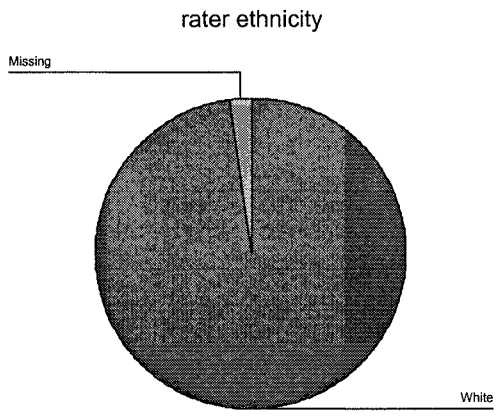
	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
EXTROV Equal variances assumed	2.717	.102	-.374	133	.709	-.3933	1.0519	-2.4739	1.6872
Equal variances not assumed			-.459	24.764	.650	-.3933	.8576	-2.1604	1.3738
NEUROT Equal variances assumed	.786	.377	-.138	133	.890	-.1620	1.1744	-2.4849	2.1608
Equal variances not assumed			-.136	20.712	.893	-.1620	1.1894	-2.6377	2.3136
LIE Equal variances assumed	3.449	.065	-.676	133	.500	-.3056	.4523	-1.2002	.5890
Equal variances not assumed			-.794	23.736	.435	-.3056	.3847	-1.1001	.4889
AGE Equal variances assumed	.792	.375	4.036	132	.000	8.1478	2.0188	4.1545	12.1411
Equal variances not assumed			3.532	19.394	.002	8.1478	2.3069	3.3260	12.9697
W.G Equal variances assumed	5.063	.026	1.084	133	.280	2.4043	2.2185	-1.9839	6.7925
Equal variances not assumed			1.502	28.760	.144	2.4043	1.6008	-.8709	5.6795
TOTALPER Equal variances assumed	12.703	.001	-1.939	118	.055	-.4667	.2407	-.9433	9.996E-03
Equal variances not assumed			-3.348	37.948	.002	-.4667	.1394	-.7488	-.1845
IMP1 Equal variances assumed	2.807	.096	1.019	132	.310	4.7557	4.6660	-4.4742	13.9855
Equal variances not assumed			1.319	26.423	.198	4.7557	3.6046	-2.6479	12.1592
REQ1 Equal variances assumed	4.388	.038	-4.076	120	.000	-15.9065	3.9025	-23.6333	-8.1798
Equal variances not assumed			-6.420	30.461	.000	-15.9065	2.4777	-20.9636	-10.8495
IMP2 Equal variances assumed	.104	.747	-.401	132	.689	-1.6620	4.1458	-9.8628	6.5387
Equal variances not assumed			-.426	21.834	.674	-1.6620	3.9041	-9.7621	6.4381
REQ2 Equal variances assumed	1.624	.205	-1.349	120	.180	-6.9595	5.1603	-17.1765	3.2575
Equal variances not assumed			-1.851	24.424	.076	-6.9595	3.7597	-14.7119	.7929
IMP3 Equal variances assumed	.373	.542	-.860	132	.391	-3.0936	3.5960	-10.2068	4.0196
Equal variances not assumed			-.898	21.559	.379	-3.0936	3.4446	-10.2457	4.0585
REQ3 Equal variances assumed	.000	.987	-.489	105	.626	-3.0484	6.2353	-15.4119	9.3151
Equal variances not assumed			-.525	18.091	.606	-3.0484	5.8054	-15.2406	9.1439
P1 Equal variances assumed	.439	.509	-1.926	131	.056	-.3317	.1722	-.6725	9.020E-03
Equal variances not assumed			-2.337	22.469	.029	-.3317	.1419	-.6257	-3.7721E-02
P2 Equal variances assumed	1.695	.195	-1.411	118	.161	-.2884	.2044	-.6931	.1163
Equal variances not assumed			-1.609	18.203	.125	-.2884	.1792	-.6646	8.777E-02
P3 Equal variances assumed	4.051	.046	-.722	126	.472	-.1518	.2103	-.5680	.2645
Equal variances not assumed			-.933	24.530	.360	-.1518	.1626	-.4871	.1835
P4 Equal variances assumed	.555	.457	-1.273	132	.205	-.2021	.1588	-.5162	.1120
Equal variances not assumed			-1.805	29.901	.081	-.2021	.1120	-.4308	2.662E-02
D1 Equal variances assumed	2.185	.142	.977	132	.330	.1644	.1682	-.1684	.4972
Equal variances not assumed			1.043	21.916	.308	.1644	.1577	-.1626	.4914
D2 Equal variances assumed	.303	.583	-1.430	132	.155	-.2353	.1646	-.5608	9.022E-02
Equal variances not assumed			-1.926	27.789	.064	-.2353	.1222	-.4857	1.508E-02
D3 Equal variances assumed	.413	.522	-1.764	131	.080	-.3286	.1863	-.6971	3.988E-02
Equal variances not assumed			-2.466	29.427	.020	-.3286	.1333	-.6010	-5.6188E-02
T1 Equal variances assumed	.873	.352	-1.216	129	.226	-.2250	.1851	-.5911	.1412
Equal variances not assumed			-1.578	26.883	.126	-.2250	.1425	-.5175	6.754E-02
T2 Equal variances assumed	.046	.830	-1.799	125	.074	-.3384	.1881	-.7107	3.387E-02
Equal variances not assumed			-2.327	24.645	.028	-.3384	.1454	-.6381	-3.8669E-02
T3 Equal variances assumed	5.085	.026	-1.110	130	.269	-.1893	.1705	-.5266	.1481
Equal variances not assumed			-1.553	29.632	.131	-.1893	.1219	-.4383	5.982E-02
T4 Equal variances assumed	.418	.519	-1.463	128	.146	-.2764	.1889	-.6502	9.736E-02
Equal variances not assumed			-2.059	30.236	.048	-.2764	.1343	-.5506	-2.2815E-03
GM1 Equal variances assumed	.243	.623	-.190	132	.850	-3.6702E-02	.1933	-.4190	.3456
Equal variances not assumed			-.200	21.671	.844	-3.6702E-02	.1838	-.4183	.3449
GM2 Equal variances assumed	18.993	.000	-2.080	132	.039	-.3107	.1493	-.6061	-1.5286E-02
Equal variances not assumed			-2.822	28.060	.009	-.3107	.1101	-.5362	-8.5211E-02
GM3 Equal variances assumed	.347	.557	-2.416	132	.017	-.4344	.1798	-.7901	-7.8705E-02
Equal variances not assumed			-3.111	26.254	.004	-.4344	.1397	-.7213	-.1475
GM4 Equal variances assumed	.191	.663	-1.942	132	.054	-.3937	.2028	-.7947	7.395E-03
Equal variances not assumed			-2.336	24.364	.028	-.3937	.1685	-.7412	-4.6132E-02
GM5 Equal variances assumed	2.646	.106	-2.886	131	.005	-.6501	.2252	-1.0957	-.2045
Equal variances not assumed			-4.321	32.884	.000	-.6501	.1505	-.9563	-.3439
CP Equal variances assumed	.473	.493	.131	132	.896	.4403	3.3655	-6.2170	7.0977
Equal variances not assumed			.149	23.152	.883	.4403	2.9513	-5.6626	6.5433
WORKTIME Equal variances assumed	1.487	.225	.465	132	.642	.7064	1.5183	-2.2968	3.7097
Equal variances not			.545	23.730	.591	.7064	1.2962	-1.9705	3.3833

assumed								
PROMTNS Equal variances assumed	.534	.466	.185	132	.854	6.988E-02	.3783	-.6785	.8183
Equal variances not assumed			.200	22.174	.843	6.988E-02	.3492	-.6541	.7938
SALARY Equal variances assumed	10.768	.001	4.172	132	.000	8874.7290	2127.4658	4666.3914	13083.0666
Equal variances not assumed			7.860	56.673	.000	8874.7290	1129.1492	6613.3641	11136.0940
ABSENCE Equal variances assumed	.533	.467	-.622	132	.535	-.1116	.1795	-.4668	.2435
Equal variances not assumed			-.548	19.459	.590	-.1116	.2038	-.5375	.3142

Rater characteristics







MANAGER

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		1	.7	.7	.7
	c.t.	1	.7	.7	1.5
	c.t.b.	1	.7	.7	2.2
	c.w.	2	1.5	1.5	3.7
	d.g.	1	.7	.7	4.4
	d.k.	3	2.2	2.2	6.7
	d.s.	2	1.5	1.5	8.1
	g.e.	1	.7	.7	8.9
	g.mc.c	1	.7	.7	9.6
	g.w.	4	3.0	3.0	12.6
	j.p.	2	1.5	1.5	14.1
	p.d.	1	.7	.7	14.8
	p.h.	1	.7	.7	15.6
	r1	4	3.0	3.0	18.5
	r10-1	2	1.5	1.5	20.0
	r10	7	5.2	5.2	25.2
	r11	2	1.5	1.5	26.7
	r12	3	2.2	2.2	28.9
	r13	4	3.0	3.0	31.9
	r14	1	.7	.7	32.6
	r15	4	3.0	3.0	35.6
	r16	3	2.2	2.2	37.8
	r17	4	3.0	3.0	40.7
	r18	1	.7	.7	41.5
	r19	7	5.2	5.2	46.7
	r2	1	.7	.7	47.4
	r20	3	2.2	2.2	49.6
	r21	2	1.5	1.5	51.1
	r22	2	1.5	1.5	52.6
	r23	16	11.9	11.9	64.4
	r24	4	3.0	3.0	67.4
	r25-1	1	.7	.7	68.1
	r25	3	2.2	2.2	70.4
	r26	4	3.0	3.0	73.3
	r27	1	.7	.7	74.1
	r27	1	.7	.7	74.1
	r28	4	3.0	3.0	77.0
	r28	4	3.0	3.0	77.0
	r29	5	3.7	3.7	80.7
	r29	5	3.7	3.7	80.7
	r3	2	1.5	1.5	82.2
	r3	2	1.5	1.5	82.2
	r30	2	1.5	1.5	83.7
	r30	2	1.5	1.5	83.7
	r31	4	3.0	3.0	86.7
	r31	4	3.0	3.0	86.7
	r4	3	2.2	2.2	88.9
	r4	3	2.2	2.2	88.9
	r5	4	3.0	3.0	91.9
	r5	4	3.0	3.0	91.9
	r6	4	3.0	3.0	94.8
	r6	4	3.0	3.0	94.8
	r7	2	1.5	1.5	96.3
	r7	2	1.5	1.5	96.3
	r8	4	3.0	3.0	99.3
	r8	4	3.0	3.0	99.3
	r9	1	.7	.7	100.0
	r9	1	.7	.7	100.0
	Total	135	100.0	100.0	
	Total	135	100.0	100.0	

MANAGER

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		1	.7	.7	.7
	c.t.	1	.7	.7	1.5
	c.t.b.	1	.7	.7	2.2
	c.w.	2	1.5	1.5	3.7
	d.g.	1	.7	.7	4.4
	d.k.	3	2.2	2.2	6.7
	d.s.	2	1.5	1.5	8.1
	g.e.	1	.7	.7	8.9
	g.mc.c	1	.7	.7	9.6
	g.w.	4	3.0	3.0	12.6
	j.p.	2	1.5	1.5	14.1
	p.d.	1	.7	.7	14.8
	p.h.	1	.7	.7	15.6
	r1	4	3.0	3.0	18.5
	r10-1	2	1.5	1.5	20.0
	r10	7	5.2	5.2	25.2
	r11	2	1.5	1.5	26.7
	r12	3	2.2	2.2	28.9
	r13	4	3.0	3.0	31.9
	r14	1	.7	.7	32.6
	r15	4	3.0	3.0	35.6
	r16	3	2.2	2.2	37.8
	r17	4	3.0	3.0	40.7
	r18	1	.7	.7	41.5
	r19	7	5.2	5.2	46.7
	r2	1	.7	.7	47.4
	r20	3	2.2	2.2	49.6
	r21	2	1.5	1.5	51.1
	r21	2	1.5	1.5	51.1
	r22	2	1.5	1.5	52.6
	r22	2	1.5	1.5	52.6
	r23	16	11.9	11.9	64.4
	r23	16	11.9	11.9	64.4
	r24	4	3.0	3.0	67.4
	r24	4	3.0	3.0	67.4
	r25-1	1	.7	.7	68.1
	r25-1	1	.7	.7	68.1
	r25	3	2.2	2.2	70.4
	r25	3	2.2	2.2	70.4
	r26	4	3.0	3.0	73.3
	r26	4	3.0	3.0	73.3
	r27	1	.7	.7	74.1
	r27	1	.7	.7	74.1
	r28	4	3.0	3.0	77.0
	r28	4	3.0	3.0	77.0
	r29	5	3.7	3.7	80.7
	r29	5	3.7	3.7	80.7
	r3	2	1.5	1.5	82.2
	r3	2	1.5	1.5	82.2
	r30	2	1.5	1.5	83.7
	r30	2	1.5	1.5	83.7
	r31	4	3.0	3.0	86.7
	r31	4	3.0	3.0	86.7

r4	3	2.2	2.2	88.9
r5	4	3.0	3.0	91.9
r6	4	3.0	3.0	94.8
r7	2	1.5	1.5	96.3
r8	4	3.0	3.0	99.3
r9	1	.7	.7	100.0
Total	135	100.0	100.0	

Development of models from Managerial Sample

Items included in each model are included in appendix VI

Reliability analyses of the models

Critical Thinking Analogue model
N of Cases = 113.0 N of Items = 19 Alpha = .4704

Extroversion analogue model
N of Cases = 128.0 N of Items = 27 Alpha = .6515

Neuroticism analogue model
N of Cases = 132.0 N of Items = 21 Alpha = .6195

Construct validity of biodata analogue models: Correlation of analogue and psychometric predictors

Correlations

		WG2WI	WG2RAT	W.G
WG2WI	Pearson Correlation	1.000	.999	.219
	Sig. (2-tailed)	.	.000	.020
	N	113	113	113
WG2RAT	Pearson Correlation	.999	1.000	.202
	Sig. (2-tailed)	.000	.	.032
	N	113	113	113
W.G	Pearson Correlation	.219	.202	1.000
	Sig. (2-tailed)	.020	.032	.
	N	113	113	135

Correlations

		EX2WI	EX2RAT	EXTROV
EX2WI	Pearson Correlation	1.000	.984	.732
	Sig. (2-tailed)	.	.000	.000
	N	128	128	128
EX2RAT	Pearson Correlation	.984	1.000	.704
	Sig. (2-tailed)	.000	.	.000
	N	128	128	128
EXTROV	Pearson Correlation	.732	.704	1.000
	Sig. (2-tailed)	.000	.000	.
	N	128	128	135

Correlations

		N2WI	N2RAT	NEUROT
N2WI	Pearson Correlation	1.000	.988	.630
	Sig. (2-tailed)	.	.000	.000
	N	132	132	132
N2RAT	Pearson Correlation	.988	1.000	.617
	Sig. (2-tailed)	.000	.	.000
	N	132	132	132
NEUROT	Pearson Correlation	.630	.617	1.000
	Sig. (2-tailed)	.000	.000	.
	N	132	132	135

Incremental validity analysis

Step 1: Observation of correlations between psychometric predictions, biodata analogue prediction and performance outcomes

The table below presents the correlations between the Watson Glaser **critical thinking appraisal** score, the biodata analogue models and the performance outcomes of total performance rating and career progress.

NB. Career progress = number of promotions / time in the organisation * 100

Correlations

		W.G	WG2WI	WG2RAT	TOTALPER	CP
W.G	Pearson Correlation	1.000	.219*	.202*	.099	-.043
	Sig. (2-tailed)	.	.020	.032	.283	.62
	N	135	113	113	120	13
WG2WI	Pearson Correlation	.219*	1.000	.999**	-.128	-.17
	Sig. (2-tailed)	.020	.	.000	.204	.06
	N	113	113	113	100	11
WG2RAT	Pearson Correlation	.202*	.999**	1.000	-.131	-.17
	Sig. (2-tailed)	.032	.000	.	.194	.06
	N	113	113	113	100	11
TOTALPER	Pearson Correlation	.099	-.128	-.131	1.000	.00
	Sig. (2-tailed)	.283	.204	.194	.	.97
	N	120	100	100	120	11
CP	Pearson Correlation	-.043	-.173	-.173	.002	1.00
	Sig. (2-tailed)	.621	.068	.068	.979	
	N	134	112	112	119	13

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

The table below presents the correlations between **extroversion** score, the biodata analogue models and the performance outcomes of total performance rating and career progress.

Correlations

		EXTROV	EX2WI	EX2RAT	TOTALPER	CP
EXTROV	Pearson Correlation	1.000	.732**	.704**	-.011	.038
	Sig. (2-tailed)	.	.000	.000	.902	.664
	N	135	128	128	120	134
EX2WI	Pearson Correlation	.732**	1.000	.984**	.023	.152
	Sig. (2-tailed)	.000	.	.000	.811	.087
	N	128	128	128	113	127
EX2RAT	Pearson Correlation	.704**	.984**	1.000	.043	.156
	Sig. (2-tailed)	.000	.000	.	.654	.080
	N	128	128	128	113	127
TOTALPER	Pearson Correlation	-.011	.023	.043	1.000	.002
	Sig. (2-tailed)	.902	.811	.654	.	.979
	N	120	113	113	120	119
CP	Pearson Correlation	.038	.152	.156	.002	1.000
	Sig. (2-tailed)	.664	.087	.080	.979	
	N	134	127	127	119	134

**. Correlation is significant at the 0.01 level (2-tailed).

The table below presents the correlations between **neuroticism score**, the biodata analogue models and the performance outcomes of total performance rating and career progress.

Correlations

		NEUROT	N2WI	N2RAT	TOTALPER	CP
NEUROT	Pearson Correlation	1.000	.630**	.617**	-.069	-.039
	Sig. (2-tailed)	.	.000	.000	.454	.656
	N	135	132	132	120	134
N2WI	Pearson Correlation	.630**	1.000	.988**	-.086	-.100
	Sig. (2-tailed)	.000	.	.000	.354	.258
	N	132	132	132	117	131
N2RAT	Pearson Correlation	.617**	.988**	1.000	-.073	-.080
	Sig. (2-tailed)	.000	.000	.	.432	.361
	N	132	132	132	117	131
TOTALPER	Pearson Correlation	-.069	-.086	-.073	1.000	.002
	Sig. (2-tailed)	.454	.354	.432	.	.979
	N	120	117	117	120	119
CP	Pearson Correlation	-.039	-.100	-.080	.002	1.000
	Sig. (2-tailed)	.656	.258	.361	.979	
	N	134	131	131	119	134

**. Correlation is significant at the 0.01 level (2-tailed).

Building empirical 'post-hoc' CP and TP models

Significant correlations with CP
q5b -.171 (133) p=0.049
q8_1 .181 (134) p = 0.34
q49 -.208 (133) p = .016
q64 -.286 (133) p = .001
q22-3 -.171 (132) p = .050

These items created 'CP-emp' which significantly correlated with original CP at 0.216 (131) p = 0.013

performance data correlations
imp3 -.203 (133) p = .019
d1 .185 (133) p = .033
promotions .774 (134) p = .000
q2 (age) -.249 (120) .000
q44-7 .198 (134) p = 0.22 not used as only 8.1% of responses were positive
q40-3 .241 (133)p = .005 not used as only 0.7% of responses were positive
q26-5 .206 (134) p = .017 not used as only 3% of responses were positive
q27-3 .198 (134) p = .022 not used as only 2.2% of responses were positive
q28-4 .182 (134) p = .035 not used as only 3% of responses were positive

```
Reliability Coefficients
N of Cases =    133.0      N of Items =  6  Alpha =    .0553
```

Significant correlations with Total performance
Q2 (age) -.330 (120) .000
Q8-5 .181 (120) p = 0.048
Q9 .210 (119) p = 0.022
Q26-1 .231 (120) p = .011
Q11-4 -.220 (116) p = .018
Q28-2 .206 (120) p = .024

These items created 'totp-emp' which significantly correlated with original total performance at 0.407 (115)P = 0.000
Q44-1 -.189 (120) p = 0.038 not used as only 5.9% of responses were positive
Q21-1 0.204 (119) p =0.026 not used as only 5.2% of responses were positive
Q7-8 .224 (118) p = 0.015 not used as only 7.4% of responses were positive
Q11-3 .215 (116) p = .021 not used as only 4.4% of responses were positive

```
Reliability Coefficients
N of Cases =    130.0      N of Items =  6    Alpha =    .1931
```

GM3 – perceived energy

Items with significant correlations with GM3 to be included in a weighted item empirical model.

Q2 (age) -.272 (134) p = .001
Male -.206 (134) p = .017
Q8-5 .203 (134) p = .019
Q21-2 .220 (133) p = .011
Q26-1 .195 (134) p = .024
Q61-1 .199 (133) p = .022
Q61-4 .170 (133) p =.050
Q32-4 -.179 (134) p = .038
Q28-2 -.188 (134) p = .029
Q3 .186 (131) p = .033
Q5a .204 (134) p = .018
Q5b -.200 (134) p = .021
Q16 .226 (134) p =.009
Q24 -.174 (133) p =.045
Q52 .172 (134) p =.047
Q59 .243 (134) p = .005
Q60 .197 (134) p = .022
Q63 -.193 (134) p =.026
Q64 -.194 (1340 p = .025

Q32-2 .225 (134) p = .009 not used as only 9.6% of responses were positive
Q44-1 -.209 (134) p = .015 not used as only 5.9% of responses were positive
Q11-3 .197 (130) p = .025not used as only 4.6% of responses were positive

Correlation with original GM3
0. 245 (128) P = 0.005

```
Reliability Coefficients
N of Cases =    128.0     N of Items = 19     Alpha =    .4916
```

Cross validation exercise

Correlation of models developed from the pilot samples onto the managerial sample

Correlations

		MWIWG	MRATWG	W.G
MWIWG	Pearson Correlation	1.000	.999	.238
	Sig. (2-tailed)	.	.000	.012
	N	110	110	110
MRATWG	Pearson Correlation	.999	1.000	.220
	Sig. (2-tailed)	.000	.	.021
	N	110	110	110
W.G	Pearson Correlation	.238	.220	1.000
	Sig. (2-tailed)	.012	.021	.
	N	110	110	135

Correlations

		MWIEX	MRATEX	EXTROV
MWIEX	Pearson Correlation	1.000	.983	.716
	Sig. (2-tailed)	.	.000	.000
	N	126	126	126
MRATEX	Pearson Correlation	.983	1.000	.673
	Sig. (2-tailed)	.000	.	.000
	N	126	126	126
EXTROV	Pearson Correlation	.716	.673	1.000
	Sig. (2-tailed)	.000	.000	.
	N	126	126	135

Correlations

		MWIN	MRATN	NEUROT
MWIN	Pearson Correlation	1.000	.991	.636
	Sig. (2-tailed)	.	.000	.000
	N	129	129	129
MRATN	Pearson Correlation	.991	1.000	.622
	Sig. (2-tailed)	.000	.	.000
	N	129	129	129
NEUROT	Pearson Correlation	.636	.622	1.000
	Sig. (2-tailed)	.000	.000	.
	N	129	129	135

Correlations of Wilkinson's (1995) models onto the current sample of managerial job incumbents

Watson Glaser weighted item and rational models

Correlations

		LWIWG	LRATWG	W.G
LWIWG	Pearson Correlation	1.000	.980	.452
	Sig. (2-tailed)	.	.000	.000
	N	116	116	116
LRATWG	Pearson Correlation	.980	1.000	.461
	Sig. (2-tailed)	.000	.	.000
	N	116	116	116
W.G	Pearson Correlation	.452	.461	1.000
	Sig. (2-tailed)	.000	.000	.
	N	116	116	135

Extroversion weighted item and rational models

Correlations

		LWIEX	LRATEX	EXTROV
LWIEX	Pearson Correlation	1.000	.969	.567
	Sig. (2-tailed)	.	.000	.000
	N	132	132	132
LRATEX	Pearson Correlation	.969	1.000	.507
	Sig. (2-tailed)	.000	.	.000
	N	132	132	132
EXTROV	Pearson Correlation	.567	.507	1.000
	Sig. (2-tailed)	.000	.000	.
	N	132	132	135